II

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/82, 9/12, C07K 14/415, A01H 5/00 — A1

(11) International Publication Number: WO 97/43427

(43) International Publication Date: 20 November 1997 (20.11.97)

(21) International Application Number: PCT/EP97/02443

(22) International Filing Date: 13 May 1997 (13.05.97)

(30) Priority Data:
9610044.1 — 14 May 1996 (14.05.96) — GB

(71) Applicant (*for all designated States except US*): NOVARTIS AG [CH/CH]; Scharzwaldallee 215, CH-4058 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): DE VRIES, Sape, Cornelis [NL/NL]; Roghorst 192, NL-6708 KS Wageningen (NL). SCHMIDT, Eduard, Daniel, Leendert [NL/NL]; Callunastraat 25, NL-6813 ET Arnheim (NL). VAN HOLST, Gerrit, Jan [NL/NL]; De Gouw 8, NL-1602 DN Enkhuizen (NL). HECHT, Valerie, France, Gabrielle [NL/NL]; Kees Muldenweg 25, NL-6707 HA Wageningen (NL).

(74) Agent: ROTH, Bernhard, M.; Novartis AG, Patent- und Markenabteilung, Klybeckstrasse 141, CH-4002 Basel (CH).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PRODUCTION OF APOMICTIC SEED

(57) Abstract

The present invention provides, *inter alia*, a method of producing apomictic seeds comprising the steps of: (i) transforming plant material with a nucleotide sequence encoding a protein the presence of which in a cell, or membrane thereof, renders said cell embryogenic, (ii) regenerating the thus transformed material into plants, or carpel containing parts thereof, and (iii) expressing the sequence in the vicinity of the embryo sac. The protein may be a leucine repeat rich receptor kinase which preferably is modified to the extent that the ligand binding domain is deleted or functionally inactivated.

close prior art but
Seq is very different
from applicants

potential

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Production of Apomictic Seed

The present invention relates to the production of genetically transformed plants. In particular the invention relates *inter alia* to a process for inducing apomixis, to the apomictic seeds which result from the process, and to the plants and progeny thereof which result from the germination of such seeds.

Apomixis, which is vegetative (non-sexual) reproduction through seeds, is a genetically controlled reproductive mechanism found in some polyploid non-cultivated species. The process is classified as gametophytic or non-gametophytic. In gametophytic apomixis - of which there are two types (apospory and diplospory), multiple embryo sacs which typically lack antipodal nuclei are formed, or else megasporogenesis in the embryo sac takes place. In adventitious embryony (non-gametophytic apomixis), a somatic embryo develops directly from the cells of the embryo sac, ovary wall or integuments. In adventitious embryony, somatic embryos from surrounding cells invade the sexual ovary, one of the somatic embryos out-competes the other somatic embryos and the sexual embryo and utilizes the produced endosperm.

Were apomixis to be a controllable and reproducible phenomenon it would provide many advantages in plant improvement and cultivar development in the case that sexual plants are available as crosses with the apomictic plant.

For example, apomixis would provide for true-breeding, seed propagated hybrids. Moreover, apomixis could shorten and simplify the breeding process so that selfing and progeny testing to produce and/or stabilize a desirable gene combination could be eliminated. Apomixis would provide for the use as cultivars of genotypes with unique gene combinations since apomictic genotypes breed true irrespective of heterozygosity. Genes or groups of genes could thus be "pyramided and "fixed" in super genotypes. Every superior apomictic genotype from a sexual-apomictic cross would have the potential to be a cultivar. Apomixis would allow plant breeders to develop cultivars with specific stable traits for such characters as height, seed and forage quality and maturity. Breeders would not be limited in their commercial production of hybrids by (i) a cytoplasmic-nuclear interaction to produce male sterile female parents or (ii) the fertility restoring

capacity of a pollinator. Almost all cross-compatible germplasm could be a potential parent to produce apomictic hybrids.

Finally, apomixis would simplify commercial hybrid seed production. In particular, (i) the need for physical isolation of commercial hybrid production fields would be eliminated; (ii) all available land could be used to increase hybrid seed instead of dividing space between pollinators and male sterile lines; and (iii) the need to maintain parental line seed stocks would be eliminated.

The potential benefits to accrue from the production of seed *via* apomixis are presently unrealized, to a large extent because of the problem of engineering apomictic capacity into plants of interest. The present invention provides a solution to that problem in that it provides the means for obtaining plants which exhibit the adventitious embryony type of apomixis.

According to the present invention there is provided a method of producing apomictic seeds comprising the steps of:

(i) transforming plant material with a nucleotide sequence encoding a protein the presence of which in an active form in a cell, or membrane thereof, renders said cell embryogenic,

(ii) regenerating the thus transformed material into plants, or carpel-containing parts thereof, and

(iii) expressing the sequence in the vicinity of the embryo sac.

By "vicinity of the embryo sac" is meant in one or more of the following: carpel, integuments, ovule, ovule premordium, ovary wall, chalaza, nucellus, funicle and placenta. The skilled man will recognize that the term "integuments" also includes those tissues, such as endothelium, which are derived therefrom. By "embryogenic" is meant the capability of cells to develop into an embryo under permissive conditions. It will be appreciated that the term "in an active form" includes proteins which are truncated or otherwise mutated with the proviso that they initiate or amplify embryogenesis whether or not in doing this they interact with the signal transduction components that they otherwise would in the tissues in which they are normally present.

The term "plant material" includes protoplasts, isolated plant cells (such as stomatal guard cells) possessing a cell wall, pollen, whole tissues such as emerged radicle, stem, leaf, petal,

hypocotyl section, apical meristem, ovaries, zygotic embryo *per se*, roots, vascular bundle, pericycle, anther filament, somatic embryos and the like.

A further embodiment of the invention relates to a DNA molecule comprising a nucleotide sequence encoding a protein the presence of which in an active form in a cell, or membrane thereof, renders said cell embryogenic.

The said nucleotide sequence may be introduced into the plant material, *inter alia, via* a bacterial or viral vector, by micro-injection, by co-incubation of the plant material and sequence in the presence of a high molecular weight glycol or by coating of the sequence onto the surface of a biologically inert particle which is then introduced into the material.

Expression of the sequence may yield a protein kinase capable of spanning a plant cell membrane. Typically the kinase may be a leucine rich repeat receptor like kinase which has the capacity to auto-phosphorylate. The skilled man will recognize what is meant by the term "leucine rich repeat receptor like kinase". Examples of such proteins include *Arabidopsis* RLK5 (Walker, 1993), *Arabidopsis* RPS2 (Bent *et al.* 1994), Tomato *CF-9* gene product (Jones *et al.* 1994), Tomato N (Whitham *et al.* 1994), *Petunia* PRK1 (Mu *et al.* 1994), the product of the *Drosophila Toll* gene (Hashimoto *et al.* 1988), the protein kinase encoded by the rice *OsPK10* gene (Zhao *et al.* 1994), the translation product of the rice EST clone ric2976 and the product of the *Drosophila Pelle* gene (Shelton and Wasserman, 1993). Still further examples of such proteins include the TMK1, Clavatal, Erecta, and TMKL1 gene products from *Arabidopsis*, the Flightless-1 gene product from *Drosophila*, the TrkC gene product from pig, the rat LhCG receptor and FSH receptor, the dog TSH receptor, and the human Trk receptor kinase. The protein may comprise a ligand binding domain, a proline box, a transmembrane domain, a kinase domain and a protein binding domain. In many receptor kinases the extracellular (ligand binding) domain serves as an inhibitor of the kinase domain in the ligand-free state. This arrest is removed after binding of the ligand. Accordingly, in one embodiment of the invention the protein either lacks a ligand binding domain or the domain is functionally inactivated so that the kinase domain can be constitutively active in the absence of an activating signal (ligand). Whether or not the protein possesses a ligand binding domain - functional or otherwise, once expressed and incorporated into the plant cell membrane the protein binding domain is preferably located intra-cellularly.

In a preferred embodiment of the method, the said sequence further encodes a cell membrane targeting sequence. The sequence may be that which is depicted in SEQ ID Nos. 1, 2, 20, or 32, or it may be similar in that it is complementary to a sequence which hybridizes under stringent conditions with the said sequences and which encodes a membrane bound protein having kinase activity. By "similar" is meant a sequence which is complementary to a test sequence which is capable of hybridizing to the inventive sequence. When the test and inventive sequences are double stranded the nucleic acid constituting the test sequence preferably has a TM within 20°C of that of the inventive sequence. In the case that the test and inventive sequences are mixed together and denatured simultaneously, the TM values of the sequences are preferably within 10°C of each other. More preferably the hybridization is performed under stringent conditions, with either the test or inventive DNA preferably being supported. Thus either a denatured test or inventive sequence is preferably first bound to a support and hybridization is effected for a specified period of time at a temperature of between 50 and 70°C in double strength citrate buffered saline (SSC) containing 0.1%SDS followed by rinsing of the support at the same temperature but with a buffer having a reduced SSC concentration. Depending upon the degree of stringency required, and thus the degree of similarity of the sequences, at a particular temperature, - such as 60°C, for example - such reduced concentration buffers are typically single strength SSC containing 0.1%SDS, half strength SSC containing 0.1%SDS and one tenth strength SSC containing 0.1%SDS. Sequences having the greatest degree of similarity are those the hybridization of which is least affected by washing in buffers of reduced concentration. It is most preferred that the test and inventive sequences are so similar that the hybridization between them is substantially unaffected by washing or incubation in one tenth strength sodium citrate buffer containing 0.1%SDS.

Accordingly, further comprised by the present invention is a DNA sequence as depicted in SEQ ID NOS: 22, 24, 26, 28 and 30 or a sequence which is complementary to one which hybridizes under stringent conditions with the said sequences and which encodes a membrane bound protein having kinase activity.

The sequence may be modified in that known mRNA instability motifs or polyadenylation signals may be removed and/or codons which are preferred by the plant into which the sequence is to be inserted may be used so that expression of the thus modified sequence in the said plant may yield substantially similar protein to that obtained by expression of the unmodified sequence in the organism in which the protein is endogenous.

In order to obtain expression of the sequence in the regenerated plant (and in particular the carpel thereof) in a tissue specific manner the sequence is preferably under expression control of an inducible or developmentally regulated promoter, typically one of the following: a promoter which regulates expression of SERK genes *in planta*, the *Arabidopsis* ANT gene promoter, the promoter of the O126 gene from *Phalaenopsis*, the carrot chitinase DcEP3-1 gene promoter, the *Arabidopsis* AtChitIV gene promoter, the *Arabidopsis* LTP-1 gene promoter, the *Arabidopsis* bel-1 gene promoter, the petunia fbp-7 gene promoter, the *Arabidopsis* AtDMC1 promoter, the pTA7001 inducible promoter. The DcEP3-1 gene is expressed transiently during inner integument degradation and later in cells that line the inner part of the developing endosperm. The AtChiIV gene is transiently expressed in the micropylar endosperm up to cellularisation. The LTP-1 promoter is active in the epidermis of the developing nucellus, both integuments, seed coat and early embryo. The bel-1 gene is expressed in the developing inner integument and the fbb-7 promoter is active during embryo sac development. The *Arabidopsis* ANT gene is expressed during integument development, and the O126 gene from *Phalaenopsis* is expressed in the mature ovule.

It is most preferred that the sequence is expressed in the somatic cells of the embryo sac, ovary wall, nucellus, or integuments.

The endosperm within the apomictic seed results from fusion of polar nuclei within the embryo sac with a pollen-derived male gamete nucleus. It is preferred that the sequence encoding the protein is expressed prior to fusion of the polar nuclei with the male gamete nucleus.

The invention further includes a DNA, but preferably a recombinant DNA, comprising a sequence encoding a protein the presence of which in an active form in a cell, or membrane thereof, renders said cell embryogenic. Preferred is a DNA encoding a protein which is a leucine rich repeat receptor like kinase and comprises a ligand binding domain, a proline box, a transmembrane domain, a kinase domain and a protein binding domain, the ligand binding domain optionally being absent or functionally inactive.
In particular, the invention embodies a DNA comprising a DNA sequence encoding a N-terminal protein fragment having the following amino acid sequence: Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val Thr Cys Asn.

A specific embodiment of the invention relates to a DNA comprising a DNA sequence encoding a protein having the sequence depicted in SEQ ID Nos. 3 or 21, or a protein substantially similar thereto which is capable of being membrane bound and which has kinase activity. By substantially similar is meant a pure protein having an amino acid sequence which is at least 90% similar to the sequence of the proteins depicted in SEQ ID No 3 below. In the context of the present invention, two amino acid sequences with at least 90% similarity to each other have at least 90% identical or conservatively replaced amino acid residues in a like position when aligned optimally allowing for up to 8 gaps with the *proviso* that in respect of each gap a total not more than 4 amino acid residues is affected. For the purpose of the present invention conservative replacements may be made between amino acids within the following groups:

(i) Serine and Threonine;

(ii) Glutamic acid and Aspartic acid;

(iii) Arginine and Lysine;

(iv) Asparagine and Glutamine;

(v) Isoleucine, Leucine, Valine and Methionine;

(vi) Phenylalanine, Tyrosine and Tryptophan

(vii) Alanine and Glycine

In addition, non-conservative replacements may also occur at a low frequency. Accordingly, the invention futher embodies a DNA comprising a DNA sequence encoding a N-terminal protein fragment having the following amino acid sequence: Val Xaa Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val Thr Cys Asn, with Xaa being a variable amino acid, but preferably Leu or Val.

Especially preferred within the scope of the invention is a DNA comprising a DNA sequence encoding a N-terminal protein fragment having the following amino acid sequence: Val Xaa Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val Thr Cys Asn Xab Xac Xad Xae Val Xaf Arg Val Asp Leu Gly Asn Xag Xah Leu Ser Gly His Leu Xai Pro Glu Leu Gly Xaj Leu Xak Xal Leu Gln, with Xaa to Xak representing variable amino acids, but preferably

Xaa = Leu or Val

Xab = Asn or Gln

Xac = Glu or Asp or His

Xad = Asn or His

Xae = Ser or Arg or Gln

Xaf = Ile or Thr
Xag = Ala or Ser
Xah = Glu or Asn
Xai = Val or Ala
Xaj = Val or Lys
Xak = Lys or Glu
Xal = Asn or His

It is preferred that the DNA further encodes a cell membrane targeting sequence, and that the protein encoding region is under expression control of a developmentally regulated or inducible promoter, such as, for example, a promoter which regulates expression of SERK genes *in planta*, the carrot chitinase DcEP3-1 gene promoter, the *Arabidopsis* AtChitIV gene promoter, the *Arabidopsis* LTP-1 gene promoter, the *Arabidopsis* bel-1 gene promoter, the petunia fbp-7 gene promoter, the *Arabidopsis* ANT gene promoter, or the promoter of the O126 gene from *Phalaenopsis;* the *Arabidopsis* AtDMC1 promoter, or the pTA7001 inducible promoter.

Particularly preferred embodiments of the said DNA include those depicted in SEQ ID Nos. 1, 2, 20 or 32, or those which are complementary to one which hybridizes under stringent conditions with the said sequences and which encode a membrane bound protein having kinase activity. As indicated above, the DNA may be modified in that known mRNA instability motifs or polyadenylation signals may be removed and/or codons which are preferred by the plant into which the DNA is to be inserted may be used so that expression of the thus modified DNA in the said plant may yield substantially similar protein to that obtained by expression of the unmodified DNA in the organism in which the protein is endogenous.

The invention still further includes a vector which contains DNA as indicated in the three immediately preceding paragraphs, plants transformed with the recombinant DNA or vector, and the progeny of such plants which contain the DNA stably incorporated, and/or the apomictic seeds of such plants or such progeny.

The recombinant DNA molecules of the invention can be introduced into the plant cell in a number of art-recognized ways. Those skilled in the art will appreciate that the choice of method might depend on the type of plant, i.e. monocot or dicot, targeted for transformation. Suitable methods of transforming plant cells include microinjection

(Crossway *et al., BioTechniques 4*:320-334 (1986)), electroporation (Riggs *et al, Proc. Natl. Acad. Sci. USA 83*:5602-5606 (1986), *Agrobacterium* mediated transformation (Hinchee *et al., Biotechnology 6*:915-921 (1988)), direct gene transfer (Paszkowski *et al., EMBO J. 3*:2717-2722 (1984)), ballistic particle acceleration using devices available from Agracetus, Inc., Madison, Wisconsin and Dupont, Inc., Wilmington, Delaware (*see, for example,* Sanford *et al.*, U.S. Patent 4,945,050; and McCabe *et al., Biotechnology 6*:923-926 (1988)), and protoplast transformation/regeneration methods (*see* U.S. Patent No. 5,350,689 issued Sept. 27, 1994 to Ciba-Geigy Corp.). Also see, Weissinger *et al., Annual Rev. Genet. 22*:421-477 (1988); Sanford *et al., Particulate Science and Technology 5*:27-37 (1987)(onion); Christou *et al., Plant Physiol. 87*:671-674 (1988)(soybean); McCabe *et al., Bio/Technology 6*:923-926 (1988)(soybean); Datta *et al., Bio/Technology 8*:736-740 (1990)(rice); Klein *et al., Proc. Natl. Acad. Sci. USA, 85*:4305-4309 (1988)(maize); Klein *et al., Bio/Technology 6*:559-563 (1988)(maize); Klein *et al., Plant Physiol. 91*:440-444 (1988)(maize); Fromm *et al., Bio/Technology 8*:833-839 (1990); and Gordon-Kamm *et al., Plant Cell 2*:603-618 (1990)(maize).

Comprised within the scope of the present invention are transgenic plants, in particular transgenic fertile plants transformed by means of the aforedescribed processes and their asexual and/or sexual progeny, which still contain the DNA stably incorporated, and/or the apomictic seeds of such plants or such progeny.

The transgenic plant according to the invention may be a dicotyledonous or a monocotyledonous plant. Such plants include field crops, vegetables and fruits including tomato, pepper, melon, lettuce, cauliflower, broccoli, cabbage, brussels sprout, sugar beet, corn, sweetcorn, onion, carrot, leek, cucumber, tobacco, alfalfa, aubergine, beet, broad bean, celery, chicory, cow pea, endive, gourd, groundnut, papaya, pea, peanut, pineapple, potato, safflower, snap bean, soybean, spinach, squashes, sunflower, sorghum, water-melon, and the like; and ornamental crops including Impatiens, Begonia, Petunia, Pelargonium, Viola, Cyclamen, Verbena, Vinca, Tagetes, Primula, Saint Paulia, Ageratum, Amaranthus, Anthirrhinum, Aquilegia, Chrysanthemum, Cineraria, Clover, Cosmo, Cowpea, Dahlia, Datura, Delphinium, Gerbera, Gladiolus, Gloxinia, Hippeastrum, Mesembryanthemum, Salpiglossis, Zinnia, and the like. In a preferred embodiment, the DNA is expressed in "seed crops" such as corn, sweet corn and peas etc. in such a way that the apomictic seed which results from such expression is not physically mutated or otherwise damaged in comparison with seed

from untransformed like crops. Preferred are monocotyledonous plants of the *Graminaceae* family involving *Lolium*, *Zea*, *Triticum*, *Triticale*, *Sorghum*, *Saccharum*, *Bromus*, *Oryzae*, *Avena*, *Hordeum*, *Secale* and *Setaria* plants.

More preferred are transgenic maize, wheat, barley, sorghum, rye, oats, turf and forage grasses, millet and rice. Especially preferred are maize, wheat, sorghum, rye, oats, turf grasses and rice.

Among the dicotyledonous plants Arabidopsis, soybean, cotton, sugar beet, sugar cane, oilseed rape, tobacco and sunflower are more preferred herein. Especially preferred are soybean, cotton, tobacco, sugar beet and oilseed rape.

The expression 'progeny' is understood to embrace both, "asexually" and "sexually" generated progeny of transgenic plants. This definition is also meant to include all mutants and variants obtainable by means of known processes, such as for example cell fusion or mutant selection and which still exhibit the characteristic properties of the initial transformed plant, together with all crossing and fusion products of the transformed plant material. This also includes progeny plants that result from a backcrossing, as long as the said progeny plants still contain the DNA according to the invention

Another object of the invention concerns the proliferation material of transgenic plants.

The proliferation material of transgenic plants is defined relative to the invention as any plant material that may be propagated sexually or asexually *in vivo* or *in vitro*. Particularly preferred within the scope of the present invention are protoplasts, cells, calli, tissues, organs, seeds, embryos, pollen, egg cells, zygotes, together with any other propagating material obtained from transgenic plants.

Parts of plants, such as for example flowers, stems, fruits, leaves, roots originating in transgenic plants or their progeny previously transformed by means of the process of the invention and therefore consisting at least in part of transgenic cells, are also an object of the present invention. Espeically preferred a apomictic seeds.

A further object of the invention is a method of producing apomictic seeds, but preferably seeds that are of the adventitious embryony type, comprising the steps of:

(i) transforming plant material with a nucleotide sequence encoding a protein the presence of which in an active form in a cell, or membrane thereof, renders said cell

embryogenic, but preferably a protein which is a protein kinase capable of spanning a plant cell membrane and capable of autophosphorylation.

(ii) regenerating the thus transformed material into plants, or carpel-containing parts thereof, and

(iii) expressing the sequence in the vicinity of the embryo sac.

The kinase protein being expressed by the DNA according to the invention is preferably a leucine rich repeat receptor like kinase and comprises a ligand binding domain, a proline box, a transmembrane domain, a kinase domain and a protein binding domain. In a specific embodiment of the invention, the said kinase protein may lack a functional ligand binding domain but comprises a proline box, a transmembrane domain, a kinase domain and a protein binding domain.

The genetic properties engineered into the transgenic seeds and plants described above are passed on by sexual reproduction or vegetative growth and can thus be maintained and propagated in progeny plants. Generally said maintenance and propagation make use of known agricultural methods developed to fit specific purposes such as tilling, sowing or harvesting. Specialized processes such as hydroponics or greenhouse technologies can also be applied. As the growing crop is vulnerable to attack and damages caused by insects or infections as well as to competition by weed plants, measures are undertaken to control weeds, plant diseases, insects, nematodes, and other adverse conditions to improve yield. These include mechanical measures such a tillage of the soil or removal of weeds and infected plants, as well as the application of agrochemicals such as herbicides, fungicides, gametocides, nematicides, growth regulants, ripening agents and insecticides.

Use of the advantageous genetic properties of the transgenic plants and seeds according to the invention can further be made in plant breeding which aims at the development of plants with improved properties such as tolerance of pests, herbicides, or stress, improved nutritional value, increased yield, or improved structure causing less loss from lodging or shattering. The various breeding steps are characterized by well-defined human intervention such as selecting the lines to be crossed, directing pollination of the parental lines, or selecting appropriate progeny plants. Depending on the desired properties different breeding measures are taken. The relevant techniques are well known in the art and include but are not limited to hybridization, inbreeding, backcross breeding, multiline breeding,

variety blend, interspecific hybridization, aneuploid techniques, etc. Hybridization techniques also include the sterilization of plants to yield male or female sterile plants by mechanical, chemical or biochemical means. Cross pollination of a male sterile plant with pollen of a different line assures that the genome of the male sterile but female fertile plant will uniformly obtain properties of both parental lines. Thus, the transgenic seeds and plants according to the invention can be used for the breeding of improved plant lines which for example increase the effectiveness of conventional methods such as herbicide or pestidice treatment or allow to dispense with said methods due to their modified genetic properties. Alternatively new crops with improved stress tolerance can be obtained which, due to their optimized genetic "equipment", yield harvested product of better quality than products which were not able to tolerate comparable adverse developmental conditions.

In seeds production germination quality and uniformity of seeds are essential product characteristics, whereas germination quality and uniformity of seeds harvested and sold by the farmer is not important. As it is difficult to keep a crop free from other crop and weed seeds, to control seedborne diseases, and to produce seed with good germination, fairly extensive and well-defined seed production practices have been developed by seed producers, who are experienced in the art of growing, conditioning and marketing of pure seed. Thus, it is common practice for the farmer to buy certified seed meeting specific quality standards instead of using seed harvested from his own crop. Propagation material to be used as seeds is customarily treated with a protectant coating comprising herbicides, insecticides, fungicides, bactericides, nematicides, molluscicides or mixtures thereof. Customarily used protectant coatings comprise compounds such as captan, carboxin, thiram (TMTD®), methalaxyl (Apron®), and pirimiphos-methyl (Actellic®). If desired these compounds are formulated together with further carriers, surfactants or application-promoting adjuvants customarily employed in the art of formulation to provide protection against damage caused by bacterial, fungal or animal pests. The protectant coatings may be applied by impregnating propagation material with a liquid formulation or by coating with a combined wet or dry formulation. Other methods of application are also possible such as treatment directed at the buds or the fruit.

It is thus a further object of the present invention to provide plant propagation material for cultivated plants, but especially plant seed that is treated with an seed protectant coating customarily used in seed treatment.

It is a further aspect of the present invention to provide new agricultural methods such as the methods examplified above which are characterized by the use of transgenic plants, transgenic plant material, or transgenic seed according to the present invention.

To breed progeny from plants transformed according to the method of the present invention, a method such as that which follows may be used: plants produced as described in the examples set forth below are grown in pots in a greenhouse or in soil, as is known in the art, and permitted to flower. Pollen is obtained from the mature stamens and used to pollinate the pistils of the same plant, sibling plants, or any desirable plant. Similarly, the pistils developing on the transformed plant may be pollinated by pollen obtained from the same plant, sibling plants, or any desirable plant. Transformed progeny obtained by this method may be distinguished from non-transformed progeny by the presence of the introduced gene(s) and/or accompanying DNA (genotype), or the phenotype conferred. The transformed progeny may similarly be selfed or crossed to other plants, as is normally done with any plant carrying a desirable trait. Similarly, tobacco or other transformed plants produced by this method may be selfed or crossed as is known in the art in order to produce progeny with desired characteristics. Similarly, other transgenic organisms produced by a combination of the methods known in the art and this invention may be bred as is known in the art in order to produce progeny with desired characteristics.

Further comprised by the invention is a method of obtaining embryogenic cells in plant material, comprising transforming the material with a recombinant DNA sequence or a vector according to the invention, expressing the sequence in the material or derivatives thereof and subjecting the said material or derivatives to a compound which acts as a ligand for the gene product of the said sequence.

The invention further relates to a method of generating somatic embryos under *in vitro* conditions wherein the SERK protein is overexpressed ectopically.

The invention still further includes the use of the said DNA in the manufacture of apomictic seeds, in which use the sequence is expressed in the vicinity of the embryo sac.

In a specific embodiment of the invention the SERK gene may be expressed in transgenic plants such as, for example, an *Arabidopsis* plant, under the control of plant expression signals, particularly a promoter which regulates expression of SERK genes *in planta*, but preferably a

developmentally regulated or inducible promoter such as, for example, the carrot chitinase DcEP3-1 gene promoter, the *Arabidopsis* AtChitIV gene promoter, the *Arabidopsis* LTP-1 gene promoter, the *Arabidopsis* bel-1 gene promoter, the petunia fbp-7 gene promoter, the *Arabidopsis* ANT gene promoter, or the promoter of the O126 gene from *Phalaenopsis;* the *Arabidopsis* AtDMC1 promoter, or the pTA7001 inducible promoter.

The promoters of the DcEP3-1 and the AtChit IV genes may be cloned and characterized by standard procedures. The DcSERK coding sequence (SEQ ID No. 2) is cloned behind the DcEP3-1, the AtChit IV or the AtLTP-1 promoters and transformed into *Arabidopsis*. The ligation is performed in such a way that the promoter is operably linked to the sequence to be transcribed. This construct, which also contains known marker genes providing for selection of transformed material, is inserted into the T-DNA region of a binary vector such as pBIN19 and transformed into *Arabidopsis*. *Agrobacterium*-mediated transformation into *Arabidopsis* is performed by the vacuum infiltration or root transformation procedures known to the skilled man. Transformed seeds are selected and harvested and (where possible) transformed lines are established by normal selfing. Parallel transformations with 35S promoter-SERK constructs and the entire SERK gene itself are used as controls to evaluate over-expression in many cells or only in the few cells that naturally express the SERK gene. The 35S promoter-SERK construct may give embryo formation wherever the signal that activates the SERK-mediated transduction chain is present in the plant. A testing system based on emasculation and the generation of donor plant lines for pollen carrying LTP1 promoter-GUS and SERK promoter-barnase is established.

constit

tissue specific promoter

The same constructs (35S, EP3-1, AtChitIV, AtLTP-1 and SERK promoters fused to the SERK coding sequence) are employed for transformation into several *Arabidopsis* backgrounds. These backgrounds are wild type, male sterile, fis (allelic to emb 173) and primordia timing (pt)-1 lines, or a combination of two or several of these backgrounds. The wt lines are used as a control to evaluate possible effects on normal zygotic embryogenesis, and to score for seed set without fertilization after emasculation. The ms lines are used to score directly for seed set without fertilization. The fis lines exhibit a certain degree of seed and embryo development without fertilization, so may be expected to have a natural tendency for apomictic embryogenesis, which may be enhanced by the presence of the SERK constructs. The pt-1 line has superior regenerative capabilities and has been used to initiate the first stably embryogenic *Arabidopsis* cell suspension cultures. Combinations of several of the above backgrounds are obtained by

crossing with each other and with lines containing ectopic SERK expressing constructs. Except for the ms lines, propagation can proceed by normal selfing, and analysis of apomictic traits following emasculation. A similar strategy is followed in which the ATChilV, AtLTP-1 and SERK promoters are replaced by the bel-1 and fbp-7 promoters as well by other promoters specific for components of the female gametophyte. > tissue specific

Additional constructs are generated that have constitutive receptor kinase activity. Most of the receptor kinases of the SERK type act as homodimeric receptors, requiring autophosphorylation before being able to activate downstream signal transduction cascades. In many receptor kinases the extracellular domain serves as an inhibitor of the kinase domain in the ligand-free stage. This arrest is removed after binding of the ligand (Cadena and Gill, 1992). By introduction of a SERK construct, from which the extracellular ligand-binding domain has been removed, mutant homodimeric (in cells that do not have a natural population of SERK proteins) or heterodimeric (in cells that also express the unmodified forms) proteins can be generated with a constitutively activated kinase domain. This approach, when coupled to one of the promoters active in the nucellar region, results in activation of the embryogenic pathway in the absence of the activating signal. This may be an important alternative in cases where it is necessary or desirable to have activation of the SERK pathway only dependant on specific promoter activity and independent of temporal regulation of an activating signal. Introduction of SERK constructs that result in fertilization-independent-embryogenesis (fie) are tested in other species for their effect. In order to recognize the fie phenotype, the skilled man will use appropriate male sterile backgrounds. However, pollination is often necessary for apomixis of the adventitious embryony type, in order to ensure the production of endosperm.

Whilst the present invention has been particularly described by way of the production of apomictic seed by heterologous expression of the SERK gene in the nucellar region of the carpel, the skilled man will recognize that other genes, the products of which have a similar structure/function to the SERK gene product, may likewise be expressed with similar results. Moreover, although the example illustrates apomictic seed production in *Arabidopsis*, the invention is, of course, not limited to the expression of apomictic seed-inducing genes solely in this plant. Moreover, the present disclosure also includes the possibility of expressing the SERK (or related) gene sequences in the transformed plant material in a constitutive - tissue non-specific manner (for example under transcriptional control of a CaMV35S or NOS promoter). In this case, tissue specificity is assured by the localized presence within the vicinity of the embryo

sac of the ligand of the product of the said gene. Furthermore, the SERK (or related) gene products may interact with proteins such as transcription factors which are involved in regulating embryogenesis. This interaction within tissue which has been transformed according to the present disclosure is also part of the present invention.

The skilled man who has the benefit of the present disclosure will also recognize that the SERK gene (and others as indicated in the preceding paragraph) may be transformed into plant material which may be propagated and/or differentiated and used as an explant from which somatic embryos can be obtained. Expression of such sequences in the transformed tissue (which is subjected to a ligand of the kinase gene products) substantially increases the percentage of the cells in the tissue which are competent to form somatic embryos, in comparison with the number present in non-transformed like tissue.

Plant material encompasses cells + plant

The invention will be further apparent from the following description and the associated drawings and sequence listings.

SEQ ID NO. 1 depicts the *Daucus carota* genomic clone of the putative receptor kinase (SERK) associated with the transition of competent to embryogenic cells;
SEQ ID NO. 2 depicts the cDNA of the said putative kinase;
SEQ ID NOs. 3 depicts the the predicted protein sequence of the SERK protein encoded by the DNA of SEQ ID NO:1.
SEQ ID NOs: 4-16 depict the sequences of various PCR primers; and
SEQ ID NOs. 17-19 depict specific peptides contained within the gene product of SEQ ID NO. 2.
SEQ ID NO: 20 depitcts the *Arabidopsis thaliana* partial genomic clone of the putative receptor kinase (SERK) associated with the transition of competent to embryogenic cells.
SEQ ID NO: 21 depicts the predicted protein sequence of the SERK protein encoded by the DNA of SEQ ID NO:20.
SEQ ID NOs: 22, 24, 26, 28 and 30 depict the partial DNA sequences of 5 EST clones with high homology to the SERK LRR sequences .
SEQ ID NOs. 23, 25, 27, 29 and 31 depict the predicted protein sequence of the partial DNA sequences of the 5 EST clones of SEQ ID Nos: 22, 24, 26, 28 and 30.
SEQ ID NO: 32 depicts the nuclotide sequence of the SERK cDNA from *Arabidopsis thaliana*.
SEQ ID NO: 33 depicts the predicted amino acid sequence of the SERK protein from *Arabidopsis thaliana* encoded by the DNA of SEQ ID NO: 32.

Figure 1 shows the results of an RT-PCR experiment performed on RNA extracted from the indicated tissues. 40 cycles followed by Southern blotting of the resulting bands is necessary to visualize SERK expression. Lanes include explants at day 7, treated for less (lane 1) or more (lane 2) then 3 days with 2,4-D. In the original a very faint signal is visible in lane 2, but not in lane 1. Established embryogenic cultures (lanes 4-6) but not a non-embryogenic control (lane 3) express the SERK gene. In carrot plants, no expression is detectable except for developing seeds <u>after</u> pollination (lane 7). Up to day 7 after pollination, the carrot zygote remains undivided, suggesting that the observed signal is coming only from the zygote. At day 10, the early globular and at day 20 the heart stage is reached in carrot zygotic embryogenesis. No signals are seen on Northern blots.

Figure 2A shows the results of a whole-mount *in situ* hybridization with the SERK cDNA on 7 day explants treated for 3 days with 2,4 D. Few cells on the surface of the explant express the SERK gene, and those cells that do are the ones that become embryogenic. Figure 2B shows a whole mount *in situ* hybridization on a partially dissected seed containing a globular zygotic embryo. Hybridization is visualized by DIG staining.

Figure 3 shows SERK expression in embryogenic hypocotyl cells during hormone-induced activation, determined by whole mount *in situ* hybridization . Bar: 50 mm

(A-E) Cell population generated by mechanical fragmentation of the activated hypocotyls. Only few of a certain type of cell, defined enlarged cell show SERK expression (asterisks). Small cytoplasmic cells (c), enlarging cells (eg) and large cells (l) never show SERK expression.

(F) Hypocotyl longitudinal section before hormone-induced activation. It is not possible to detect any SERK expression in any type of cell.

(G-I) Proliferating mass coming from the inner hypocotyl tissues 10 days after the beginning of the hormonal treatment (longitudinal section). In G a single enlarged cells showing SERK expression is detectable within a row of negative cells showing the same morphology. In H a single enlarged cell showing serk expression is detaching from the surface of the proliferating mass. In I a cluster of enlarged cells showing SERK expression is detectable at the surface of proliferating tissue.

(J) Proliferating mass coming from the inner tissues of the hypocotyl 10 days after the beginning of the rooting treatment (24 hours with 2,4-D followed by hormone removal). Both the root primordia and the enlarged cells detaching from the surface do not show any SERK expression.

Figure 4 shows the phenotype of Arabidopsis WS plants transformed with the 2200 bp SERK-luciferase consturct at the seedling level. Pictures were taken at 28 days after germination of T2 seeds. In plant II and III no clear shoot meristem is visible at the seedling stage, 7 days after germination. The first two leaves, if they develop at all, are needleshaped as hown on the pictures taken 28 days after germination. At this time plant I, which shows no clear phenotype, already starts flowering. Secondary shoot meristems are already developing in plant no II and will also develop later from no III. Shoot meristems, influorescences and normal flowers eventually develop on all plants.

Figure 5 shows how the 2200 bp SERK luciferase construct affects the number of developing ovules in the siliques of transformed plants.

Figure 6 shows autophosphorylation of purified SERK fusion protein *in vitro*. Lane 1: purified SERK fusion protein; Lane 2: serine phosphate; Lane 3: threonine phosphate; Lane 4: thyrosine phosphate.

The following description illustrates the isolation and cloning of the SERK gene and the production of apomictic seed by heterologous expression of the said gene in the nucellar region of the carpel so that somatic embryos form which penetrate the embryo sac and are encapsulated by the seed as it develops.

ISOLATION AND CLONING OF THE SERK GENE FROM *DAUCUS CAROTA*

Isolation of cDNA clones that are preferentially expressed in embryogenic cell cultures of carrot

In order to increase the chance of success for obtaining genes expressed in carrot suspension cells competent to form embryos, the number of embryo-forming cells as present in a series of established cell cultures was determined. A sub-population of cells that passed through a 30 mm nylon sieve was isolated from eight different cultures that ranged in age between 2 months and 4 years. In these sub 30 mm populations, the number of embryos formed from the single cells and small cell clusters was determined and expressed as a percentage of the total number of cells present at the start of embryogenesis. Sieved <30 mm cultures able to form somatic embryos with a frequency of more than 1% were then used as a source for competent cells, and cultures that produced less than 0.01% embryos were used as non-embryogenic controls. As main cloning strategies, cold plaque screening (Hodge *et al.* 1992) and differential display (dd) RT-PCR (Liang and Pardee, 1992) were used besides conventional differential screening of cDNA libraries.

Labeled probes for differential screening were obtained from RNA out of a <30 mm sieved sub-population of cells from either embryogenic or non-embryogenic cell cultures. Employing these probes in a library screen of approximately 2000 plaques yielded 26 plaques that failed to show any hybridization to either probe. These so-called cold plaques were purified and used for further analysis. From the total number of plaques that did hybridize, about 30 did so only with the probe from embryogenic cells. ddRT-PCR reactions using a combination of one anchor primer and one decamer primer were performed on mRNA isolated from three embryogenic, and three non-embryogenic suspension cultures. About 50 different ddRT-PCR fragments were obtained from each reaction. Using combinations of three different anchor and six different decamer primers, a total of approximately 1000 different cDNA fragments was visualized. Six of these PCR fragments were only found in lanes made with mRNA from <30 mm populations of cells from embryogenic cultures (Table 1) and with oligo combinations of the anchor primer (5'-TTTTTTTTTTTGC-3') and the decamer primers (5'-GGGATCTAAG-3'), (5'-ACACGTGGTC-3'), (5'-TCAGCACAGG-3'). Because differential PCR fragments often consist of several unresolved cDNA fragments (Li *et al.* 1994), cloning proved to be essential prior to undertaking further characterization of the PCR fragments obtained.

All clones obtained were subjected to a second screen, that consisted of spot-dot Northern hybridization performed under conditions of high stringency. This method, that used RNA from entire unsieved embryogenic and non-embryogenic suspension cultures, proved to be a fast and reliable additional selection method. Only one clone (22-28) of the 30 clones obtained after differential screening, proved to be restricted to embryogenic cell cultures while the majority was constitutively expressed. The 26 clones obtained from the cold plaque screening required long exposure times in the spot-dot Northern analysis. Six of these clones failed to show any hybridization signal and 19 proved to be expressed in both embryogenic and non-embryogenic cell cultures. One clone (31-50) showed low expression in all embryogenic cultures, and in one non-embryogenic culture, but not in the others. Of the six cloned fragments obtained by ddRT-PCR display, four showed hybridization more or less restricted to transcripts present in embryogenic cultures. All clones that passed through the second screening were sequenced. Two of the ddRT-PCR clones (6-8 and 7-13) were identical to the carrot Lipid Transfer Protein (LTP) gene, previously identified as a marker for embryogenic carrot cell cultures. LTP expression is restricted to embryogenic cell clusters and the protoderm of somatic and zygotic embryos from the early globular stage onwards (Sterk *et al.* 1991). Therefore, while the LTP gene is not a marker for competent cells, its appearance in the screening confirms the validity of our methods with respect to the cloning of genes expressed early during somatic embryogenesis.

cDNA clone 31-50 encodes a leucine-rich repeat containing receptor-like kinase

The mRNA corresponding to the isolated clone 31-50 had an open reading frame of 1659 nucleotides encoding a protein with a calculated Mw of 55 kDa. Because clone 31-50 is mainly expressed in embryogenic cell cultures it was renamed Somatic Embryogenesis Receptor Kinase (SERK). The SERK protein contains a N-terminal domain with a five-times repeated leucine-rich motif that is proposed to act as a protein-binding region in LRR receptor kinases (Kobe and Deisenhofer, 1994). Between the extracellular LRR domain of SERK and the membrane-spanning region is a 33 amino acid region rich in prolines (13), that is unique for the SERK protein. Of particular interest is the sequence SPPPP, that is conserved in extensins, a class of universal plant cell wall proteins (Varner and Lin, 1989). The proposed intracellular domain of the protein contains the 11 subdomains characteristic of the catalytic core of protein kinases. The core sequences HRDVKAAN and GTLGYIAPE in respectively the kinase subdomains VB and VIII suggest a function as a serine / threonine kinase (Hanks *et al.* 1988).

Another interesting feature of the intracellular part of the SERK protein is that the C-terminal 24 amino-acids resembles a single LRR. The serine and threonine residues present within the intracellular LRR sequence are surrounded by acidic residues and might be targets for the autophosphorylation of SERK, thereby regulating the ability of other proteins to interact with this receptor-kinase in a similar fashion as described for the SH2 domain of the EGF family of tyrosine receptor kinases.

Hybridization of the SERK cDNA clone to the carrot genome revealed the presence of only a single main hybridizing band after digestion with EcoR1, probably reflecting a single SERK gene in the carrot genome. This was confirmed after digestion with Ddel, an enzyme that cuts three times within the SERK gene. No signal was observed after Northern blotting of mRNA from embryogenic cell cultures and hybridization with labeled SERK probes, reflecting the low levels of transcript present in these cultures. Detection of the SERK transcript on the original spot-dot Northerns was only possible after long exposure times compared with other probes.

The ability of the SERK protein to autophosphorylate was investigated *in vitro*, using a previously described autophosphorylation assay (Mu et al. 1994), with a bacterial fusion protein that contained the complete intracellular region of the SERK protein. The bacterially expressed SERK fusion protein was able to autophosphorylate, indicating that the SERK protein is able to fulfill a role as a protein kinase *in vivo* (Heldin, 1995).

Expression of the SERK gene corresponds with the first appearance of competent cells during hypocotyl activation

When carrot hypocotyls are induced with 2,4-D, only the cells of the provascular tissue proliferate. Cells of epidermal and cortical origin merely expand, suggesting that the provascular tissue derived cells form the newly initiated suspension culture. After removal of 2,4-D, the formation of somatic embryos occurs after 2-3 weeks. Somatic embryos are preceded by embryogenic cells, that are developed in turn from competent cells. While competent and embryogenic cell formation take place in the presence of 2,4-D, it was not clear when this occurred, and which cells acquired competence. Since previous experiments (Toonen et al. 1994) revealed that cell morphology is not a good criterion, the first appearance of single competent cells was determined experimentally by semi-automatic cell tracking performed on large populations of immobilized cells. Hypocotyl explants activated with 2,4-D for seven days were mechanically fragmented and samples of the resulting population of mainly single

suspension cells were immobilized to allow recording of their development by cell tracking. In the immobilized cell populations obtained in this way all the morphologically discernible cell types were present that were also seen in the un-fragmented activated hypocotyls. Because the different cell morphologies observed during hypocotyl activation were known (Guzzo *et al.* 1995), it was possible to trace back the original position of each type of cell in the activated explant. Small cytoplasm-rich cells (16x16 mm) are the proliferating cells that surround the vascular elements. Enlarging vacuolated cells (16x40 mm) are encountered on the surface of the mass of proliferating cells and these can detach from the surface when fully enlarged (35x90 mm). Large vacuolated cells (more than 60x140 mm) are the non-proliferating remnants of the hypocotyl epidermis and cortical parenchyma. The shape of the enlarging and fully enlarged cells could change from oval to elongate or triangular. Cell tracking on a total of 24,722 cells released from seven days activated hypocotyls showed that only 20 single cells formed a somatic embryo. Because of their dependance on continued 2,4-D treatment, the embryo-forming single cells are still in the competent cell stage. All of the embryo-forming single cells belonged to the category of 3,511 enlarged cells, that contained therefore competent cells in a frequency of 0.56%. The single cell tracking experiments clearly reveal that the ability of explant cells to reinitiate cell division under the influence of 2,4-D, resulting in a population of highly cytoplasmic and rapidly proliferating cells, does have a causal relation with the ability to become embryogenic. It is also clear that only a very limited number of the cells that make up the newly initiated embryogenic suspension culture are actually competent to form embryogenic cells.

Expression of the SERK gene, determined by whole mount *in situ* hybridization on a similar population of cells as used for the cell tracking experiments, was found to be restricted to only 0.44% of the enlarged cells. Therefore, the expression of the SERK gene appears closely correlated both qualitatively and quantitatively with the presence of competent single cells.

To obtain insight into the temporal regulation of SERK expression in the course of explant activation, whole mount *in situ* hybridization was performed on entire intact or hand-sectioned explants treated for different periods with 2,4-D. Representative samples were collected at daily intervals from explants untreated and treated for three days, six days, seven days or ten days with 2,4-D before returning to B5-0. No SERK-expressing cells were ever found in explants treated for less then three days with 2,4-D. While enlarged cells became present after the first five days of culture, the first few SERK-expressing enlarged cells were found after six-seven days of culture in the presence of 2,4-D treatment. These few cells were present at the surface

of the mass of proliferating cells originating from the provascular tissue. In the hypocotyls treated for ten days with 2,4-D, the number of SERK-positive cells had increased to 3.04% and included at this stage also cells present in small clusters. No SERK transcript was ever detected in small cytoplasm-rich cells or large vacuolated cells. Hypocotyls were also treated for only one day with 2,4-D and subsequently cultured in hormone-free medium for a total of seven or ten days. Under these conditions explant cells proliferated and gave rise to roots and non-embryogenic cell cultures, while SERK expression could never be detected. The in situ hybridization results described above were obtained from a relatively small number of explants and a few hundred cells, so RT-PCR followed by Southern hybridization was performed to obtain more quantitative results. These are shown in Figure 7 and confirm the close temporal correlation between the first appearance of competent cells in explants treated for three days with 2,4-D and the expression of the SERK gene. Northern hybridization never gave any signal after hybridization with SERK cDNA probes, not even after prolonged exposure in a PhosphorImager, in line with the extremely restricted expression pattern of the SERK gene.

Expression of the SERK gene corresponds with the occurrence of competent cells in established embryogenic cell cultures

While the results described so far indicate that competent and embryogenic cell formation is restricted to a particular class of enlarged cells during explant activation, the situation in an established embryogenic cell culture is more complex. Competent single cells in such cultures do not appear to belong to one cell type in particular, but have been shown to originate from all morphologically different cell types. In cell tracking experiments, embryogenic cells, that do not require exogenous auxin treatment, were never observed to be single but consisted of clusters of at least 3-4 cells (Toonen et al. 1994). SERK expression was found in all morphologically discernible single cell types that were present in an embryogenic cell culture at a frequency between 0.1 and 0.5% depending on the cell type. In non-embryogenic cultures, SERK expressing cells were never encountered. As was observed in the activated explants, SERK expression was not restricted to single cells, but also occurred in small clusters of 2 to 16 cells. Since clusters of this size are known to consist of embryogenic cells, these data show that SERK expression is not restricted to competent single cells, but may persist in small clusters of embryogenic cells. No SERK expression was encountered during the late globular, heart and torpedo-stages of somatic embryogenesis.

The SERK gene is transiently expressed in zygotic embryogenesis

The expression of the SERK gene in carrot plants was determined by RT-PCR. The results indicate that no SERK mRNA accumulates in any of the adult plant organs nor in flowers prior to pollination. The first occasion when SERK expression can be detected is in flowers at three days after pollination (DAP), at which stage fertilization has taken place and endosperm development has commenced. SERK mRNA remains present in flowers up to twenty DAP, corresponding with the early globular stage of the zygotic embryo (Yeung et al. 1996). Whole mount *in situ* hybridization on partially dissected carrot seeds confirmed that the SERK gene was only expressed in early embryos up to the globular stage. Expression was observed in the entire embryo including the suspensor. No expression was seen in seedlings, roots, stems, leaves, developing and mature flower organs, pollen grains and stigma's before and after fertilization. Tissues in the developing seed such as seed coat, integuments, all embryo sac constituents before fertilization as well as the endosperm at all stages of development investigated did not show any SERK expression. Later stages of carrot zygotic embryos were also completely devoid of SERK mRNA. Given this pattern of expression, that is restricted to the zygotic embryo, the signal as detected by RT-PCR in flowers at 3 and 7 DAP must come from SERK mRNA as present in zygotes, because in carrot the zygote remains undivided up to one week after pollination (Yeung et al. 1996). Although SERK expression persists to slightly later stages in zygotic globular embryos when compared to the somatic ones, these results confirm the transient pattern of expression as observed for the SERK gene during somatic embryogenesis and also imply that there is a correspondence between the formation of competent cells in vitro and the formation of the zygote in vivo.

METHODS

Cell culture, hypocotyl explant induction and cell tracking

Cell cultures were derived from *Daucus carota* cv. Flakkese and maintained as previously described (De Vries *et al.* 1988a). Cell suspension cultures were maintained at high cell density in B5 medium (Gamborg *et al.* 1968) supplemented with 2 mM 2,4-D (B5-2 medium). Embryo cultures with globular, heart and torpedo-stage somatic embryos were derived from <30 mm sieved cell cultures cultured at low cell density (100 000 cells / ml) in B5 medium without 2,4-D (B5-0). For hypocotyl explant induction experiments, plantlets were obtained from seed of

Daucus carota cv. S Valery as described previously (Guzzo *et al.*, 1994). The hypocotyls of one week old plantlets were divided in segments of 3-5 mm, incubated for various periods of time in B5-2 medium and returned to B5-0 medium. Seven days after explantation and exposure to 2,4-D the hypocotyl segments were fragmented on a 170 mm sieve and the resulting cells collected to form a fine cell suspension. Immobilization of these cells in B5-0.2 medium was performed in a thin layer of phytagel (Toonen *et al.* 1994). After one week of further culture 2,4-D was removed by washing the plates with B5-0 medium. This allowed embryos to develop beyond the globular stage. Recording the development of the immobilized cells was performed with a procedure modified from the previously described by Toonen et al. (1994). The main change involved a new MicroScan program for automatic 3-axis movement to scan all cells in the phytagel (Toonen et al. 1996).

Nucleic acid isolation and analysis

RNA was isolated from cultured cells and plant tissues as described by De Vries *et al.* (1988b). Poly(A)$^+$-RNA was obtained by purification by oligo (dT) cellulose (Biolabs). For RNA gel blot analysis samples of 10 mg total RNA were electrophoresed on formamide gel, and transferred to nytran-plus membranes. For RNA spot-blot analysis 5 mg of total RNA was denatured and spotted onto nytran-plus filters using a hybridot manifold (BRL).

Genomic DNA was isolated according to Sterk et al. (1991). Samples of 10 mg genomic DNA were digested with different restriction enzymes and separated on agarose gel, and transferred to nytran-plus membrane (Schleicher & Schuell). Hybridization of RNA blots took place at 42°C in hybridization buffer containing 50% formamide, 6xSSC, 5xDenhardt, 0.5% SDS and 0.1 mg/ml salm sperm DNA. Hybridization of DNA blots was performed as previously described (Sterk et al. 1991). Following hybridization, filters were washed under stringent conditions (3x20 min in 0.1% SSC, 1% SDS, at 65°C). Filters were exposed to Kodak X-Omat AR film. The integrity and the amount of RNA on the blots was confirmed by hybridization with an 18S ribosomal RNA probe. Nucleotide sequence analysis was performed on an ABI 373A automated DNA sequencer (Applied Biosystem).

Screening procedures

Two independent cDNA libraries were constructed with equal amounts of poly(A)$^+$-RNA from total established cell cultures grown for six days in B5-2 medium, sieved <125 mm cell cultures grown for six days in B5-0 medium and sieved <30 mm cell cultures grown for six days in B5-0

medium. cDNA synthesis and cloning into the Uni-ZAP™ XR vector was performed according to the manufacturers protocol (Stratagene).

Differential screening of the cDNA libraries was performed essentially as described by Scott *et al.* (1991). RNA was isolated from either three embryogenic or three non-embryogenic cell cultures, that were grown for seven days in B5-2 after sieving through 30 mm mesh. First strand cDNA synthesis was performed on 4 mg total RNA using AMV reverse transcriptase (Gibco BRL). [^{32}P]dATP labeled probes were prepared using random prime labeling on first strand cDNA. Pooled probes from embryogenic and non-embryogenic cell populations were hybridized to two pairs of nitrocellulose filters, each containing 1000 plaques from one cDNA library. After washing for 3x20 min in 0.1% SSC, 1% SDS at 65°C, hybridization was visualized by autoradiography for two days on Kodak X-omatic film. Plaques that only showed signal with the embryogenic transcript probe were purified by two further rounds of screening.

In order to identify cDNA clones which are expressed at low levels in the <30 mm sieved cell population, cold plaque screening was performed as described by Hodge *et al.* (1992). Plaques from the differential screening that did not show any signal after seven days of autoradiography were purified by two further rounds of screening. The resulting clones were used as probes for characterization of the expression pattern of the corresponding genes.

Differential Display RT-PCR

Differential display of mRNA was performed essentially as described by Liang and Pardee (1992). cDNA synthesis took place by annealing 1 mg of total RNA in 10 ml buffer containing 200 mM KCl, 10 mM Tris-HCl (pH 8.3), and 1 mM EDTA with 100 ng of one of the following anchor primers: (5'-TTTTTTTTTTTGC-3'), (5'-TTTTTTTTTTTCTG-3'), (5'-TTTTTTTTTTTCA-3'). Annealing took place by heating the mix for 3 min. at 83°C followed by incubation for 30 min at 42°C. Annealing was followed by the addition of 15 ml pre-warmed cDNA buffer containing 16 mM $MgCl_2$, 24 mM Tris-HCl (pH 8.3), 8 mM DTT, 400 mM dNTP, and 4 Units AMV reverse transcriptase (Gibco BRL). cDNA synthesis took place at 42°C for 90 min. First strand cDNA was phenol/chlorophorm extracted and precipitated with ethanol using glycogen as a carrier. The PCR reaction was performed in a reaction volume of 20 ml containing 10% of the synthesized cDNA, 100 ng of anchor primer, 20 ng of one of the following 10-mer primers: (5'-GGGATCTAAG-3') , (5'-TCAGCACAGG-3'), (5'-GACATCGTCC-3'), (5'-CCCTACTGGT-3'), (5'-ACACGTGGTC-3'), (5'-GGTGACTGTC-3'), 2 mM dNTP, 0.5 Unit *Taq* enzyme in PCR buffer (10

mM Tris-HCl (pH 9.0), 1.5 mM $MgCl_2$, 50 mM KCl, 0.01% gelatin and 0.1% Triton X100) and 6 nM [a-^{32}P] dATP (Amersham). PCR parameters were 94°C for 30 sec, 40°C for 1 min, and 72°C for 30 sec for 40 cycles using a Cetus 9600 (Perkin-Elmer). Amplified and labeled cDNAs were separated on a 6% denaturing DNA sequencing gel. Gels were dried without fixation and bands were visualized by 16 hours of autoradiography using Kodak X-omatic film. Bands containing differentially expressed cDNA fragments of 150-450 nucleotides were cut out of the gel and DNA was extracted from the gel slices by electroelution onto DE-81 paper (Whatmann). After washing of the paper in low salt buffer (100 mM $LiCl_2$ in 10 mM TE buffer), and elution of the cDNA in high salt buffer (1 M $LiCl_2$ in 10 mM TE buffer with 20% ethanol) the cDNA was concentrated by precipitation in ethanol using glycogen as carrier. Reamplification of the cDNA fragments using the same PCR cycling parameters as described above but PCR buffer containing 2.5 mM of both the 10-mer and the anchor oligo and 100 mM dNTP. DE-81 paper allowed an efficient recovery of the DNA fragments and reamplification generated an average of 500 ng DNA after 40 cycles. Amplified PCR products were blunt-ended using the Klenow fragment of *E.coli* DNA Polymerase I (Pharmacia), purified on Sephacryl-S200 columns (Pharmacia), ligated into a SmaI linearized pBluescript vector II SK⁻ (Stratagene) and transformed into *E.coli* using electroporation.

RT-PCR

Adult plant tissues from *Daucus carota* were obtained from S&G Seeds (Enkhuizen). Controlled pollination was performed by hand. Flower tissue RNA was obtained from three compete umbels for each time-point and contained all flower organs including pollen grains. 2 mg of total RNA from adult plant tissue or cell cultures was annealed at 42°C with 50 ng oligo (5'-TCTTGGACCAGATAATTC-3') in 10 ml annealing buffer (250 mM KCl, 10 mM Tris-HCl pH 8.3, 1 mM EDTA). After 30 min. annealing, 1 unit AMV-reverse transcriptase was added in a volume of 15 ml cDNA buffer (24 mM Tris-HCl pH 8.3, 16 mM $MgCl_2$, 8 mM DTT, 0.4 mM dNTP). The reverse transcription reaction took place for 90 min. at 42°C. PCR amplification of SERK-cDNA was carried out with two specific oligos for the SERK kinase domain, (5'-CTCTGATGACTTTCCAGTC-3') and (5'-AATGGCATTTGCATGG-3'). Amplification was carried out with 30 cycles of 30 sec. at 94°C, annealing at 54°C for 30 sec. and extension at 72°C for 1 min., followed by a final extension for 10 min.at 72°C.

Whole mount *in situ* hybridization

Whole mount *in situ* hybridizations were performed essentially as previously described (Engler et al. 1994). Cell cultures and somatic embryos were immobilized on poly-L-lysine coated glasses during fixation to improve handling. Whole mount *in situ* hybridization on explants took place by embedding hypocotyls from seven-days old plantlets in 3% Seaplaque agarose (Duchefa) and processing them in Eppendorf tubes. Transverse as well as longitudinal sections were made with a vibrotome (Biorad Microcut). Sections of 50-170 mm thick were incubated in B5-2 medium for a minimum of three days to induce formation of embryo-forming cells. Optimal induction was achieved with longitudinal hypocotyl sections with a thickness of at least 90 mm. To obtain proliferating, non-embryogenic cell cultures, hypocotyl sections were exposed to 2,4-D for only 1 day, and subsequently transferred to B5-0 medium (Guzzo *et al.* 1994). Whole mount in situ hybridization on developing seeds was performed by removing the chalazal end of the seeds to allow easier probe penetration. After hybridization, the enveloping layers of integuments and endosperm were carefully removed to expose the developing embryos. In situ hybridization on sections was performed as described previously (Sterk et al. 1991) except for the use of non-radioactive probes.

All samples were fixed for 60 min. in PBS containing 70 mM EGTA, 4% paraformaldehyde, 0.25% glutaraldehyde, 0.1% Tween 20, and 10% DMSO. Samples were then washed, treated with proteinase K for 10 min, again washed and fixed a second time. Hybridization solution consisted of PBS containing 0.1% Tween 20, 330 mM NaCl, 50 mg/ml heparin, and 50% deionized formamide. Hybridization took place for 16 hours at 42°C using digoxigenin-labeled sense or antisense riboprobes (Boehringer Mannheim). After washing the cells were treated with RNaseA, and incubated with anti-digoxigenin-alkaline phosphatase conjugate (Boehringer Mannheim) which had been preabsorbed with a plant protein extract. Excess antibody was removed by washing followed by rinsing in staining buffer (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$, 1 mM levamisole) and the staining reaction was performed for 16 hours in a buffer containing NBT and BCIP. Observations were performed using a Nikon Optiphot microscope equipped with Nomarski optics.

Autophosphorylation assay

A 1.4 kB SspI cDNA fragment of the SERK cDNA encoding most of the open reading frame apart from the N-terminal three LRRs was cloned into the pGEX expression vector (Pharmacia).

A fusion protein consisting of SERK and the glutathione S-transferase gene product was synthesized by a three hours induction of transformed *E.coli* with 2 mM IPTG. Fusion protein was isolated and purified as described previously (Horn and Walker, 1994). Purified fusion protein was coupled to glutathione agarose beads (Sigma) and incubated for 20 min. at 20°C in a volume of 10 ml buffer: 50 mM Hepes (pH 7.6), 10 mM $MgCl_2$, 10 mM $MnCl_2$, 1 mM DTT, 1 mCi [γ -^{32}P] (3 000 Ci/mmol) . Excess label was removed by washing the fusion protein/glutatione agarose beads three times for 5 min. in 50 mM Tris-HCl (pH 7.3), 10 mM $MgCl_2$ at 4°C. Protein was removed from the beads by cooking in SDS-PAGE loading buffer. Equal amounts of protein were separated by SDS-PAGE and protein autophosphorylation was visualized by autoradiography.

SERK fusion proteins produced with the Baculovirus expression system.

Further fusion proteins containing the intracellular part of the *Daucus carota* SERK protein (1.0 kB HindIII / SspI fragment of the carrot SERK cDNA clone 31-50) were made using the baculorvirus vector pAcHLT.
In vitro phosphorylation studies with this purified protein showed that most if not all of the autophosphorylation of this SERK fusion protein was at threonine residues (Figure 6)

Construction of viral transfer vectors

The pAcHLT-B and pAcHLT-C baculovirus transfer vectors were used for the cloning of two cDNA fragments of the carrot SERK gene. The SspI 1.41 kB fragment of carrot DcSERK cDNA was cloned into the SmaI site of pAcHLT-B and the SspI / PvuII 1.07 kB fragment of carrot DcSERK cDNA was cloned into the SmaI site of pAcHLT-C. The first construct contains the complete C-terminal part of the DcSERK protein and from the putative extracellular region the proline-rich region and three of the lecuine-rich repeats. The second construct contains only the putative intracellular region of the DcSERK gene product. Nucleotide sequence analysis was performed in order to confirm the presence and the orientation of the DcSERK cDNA within the vector.

Transformation of insect cells

The resulting transfer vectors were used to transfect (lipofect) insect cell culture Sf21 from Spodoptera frugiperda in combination with linearized AcMNPV baculovirus DNA. Monolayers of SF21 cells were transfected in 35 mm petridishes containing 2 ml of Hink's

medium. One microgram of linearized AcMNPV baculovirus DNA (Baculogold, Invitrogen) was added to 5 microgram of pAcHLT / SERK vector construct in 25 microliter of water. Fifteen microliter of Lipofectin (BRL) was mixed with 10 microliter of water, after which the DNA solution was added. After mixing 200 microliter of Hink's medium was added to the mix and the solution was transferred to the cell monolayer, from which the medium was removed. After one hour, 500 microliter of Hink's medium was added and the cells were incubated for anotehr 3 hours. Finally, 1 ml of Hink's medium with 20% foetal bovine serum (FBS) was added and the cells were incubated for 4 days. After transfection, the viral infection could be identified by the reduced growth of cells, the swollen shape and the enlarged nucleus. After four days, infected cells were harvested and the medium containing infectious budded virus was collected and used for plaque assays and amplification of recombinant virus stocks.

Isolation of single recombinant viruses

Single recombinant virus plaques were isolated from monolayers of cells infected with a titration range of the primairy virus stock. Infections was performed in 35 mm petridishes with monolayers of cells. Virus stocks were diluted in 600 microlieter of Graces medium and added to the cell monolayer, followed by a 90 minutes incubation period at in Graces medium with 20% FBS. Afterwards, 3% Sea Plaque agarose was autoclaved, mixed with an equal amount of 2x Graces medium with 20% FBS and from the resulting agarose overlay solution 2 ml. was spread over the cell monolayers after removal of the viral inoculum. After 4 days of incubation single plaques could be visulalized and purified for further analysis.

Fusion protein production.

After determining the titer of purified recombinant viruses, monolayers of Sf21 cells in 75 cm^2 flasks were infection with a multiplicity of infection (MOI) of 10. Incubation of cells with the virus inoculum was performed for 90 min. after which 8 ml. of Hink's medium with 10%FBS was added. After 3 days of incubation, cells were harvested and washed twice with PBS. Cells were lysed for 45 min on ice in twenty volumes of 1x insect cell lysis buffer (10 mM Tris pH 7.5, 130 mM NaCl, 1% Triton, 100 mM NaF, 10 mM NaPi, 10 mM NaPPi, with proteinase inhibitors: 16 mg/l benzamidine, 10 mg/l phenanthroline, 10 mg/l aprotinin, 10 mg/l leupeptin, 10 mg/l pepstatin A, 1 mM PMSF).

The lysate was cleared by centrifugation at 10.000 g for 30 min and the supernatant was batchwise incubated in TALON resin (with high affinity for the 6xHIS tag of the recombinant fusion protein). Binding was performed by gentle agitation for 20 min. at room temp. The resin was washed three times with lysis buffer, followed by an elution step with lysis buffer with 200 mM imidazole. Purified fusion protein was collected and purifty and integrity was tested by SDS-PAGE.

Autophosphorylation assays

Protein kinase activity was determined by incubating 1 microgram of purified fusion protein for 30 min. at room temp. in a buffer containing 10 mM MgCl2, 10 mM MnCl2, 1 mM DTT and 10 μM [gamma-32]ATP (10^5 pm/pmol ATP). The autophosphorylated fusion protein was purified after SDS-PAGE from the gel in a buffer containing 50 mM NH4CO3, 0.1% SDS, 0.25% beta-mercaptoethanol. Protein was precipitated with 20 μg/ml BSA and 20% (w/v) solid trichloroacetic acid. The precipitate was collected after centrifugation, hydrolysed in 50 μl 6N HCl for 1 hour at 120 degrees Celcius. HCl was subsequently removed by lyophilization and the pellet was resuspended in a buffer consiting of 2.2% formic acid and 7.8% acetic acid. Hydrolysed protein was loaded onto cellulose thin layer chromatography plates together with control amino acid samples (phosphoserine, phosphothreonine, phosphotyrosine). Chromatography was performed in a buffer containing propionic acid: 1M ammonium hydroxide: isopropyl alcohol (90:35:35 v/v/v). After separation and drying of the plates, the separated amino-acids were visualized by spraying with 0.25% ninhydrin in aceton, followed by heating for 5 min. at 65 degrees Celcius. Plates were afterwards exposed to Phospho Imager casettes in order to detect the phospho-labeled aminoacids.

SERK antibodies

Purified fusion proteins (10 μg) were mixed in complete Freund adjuvant and injected IP into BALBc mice. After 4 weeks booster antigen was injected (10 μg purified fusion protein in imcomplete Freund adjuvant). Two weeks later a final booster was injected. One week after the final booster, serum was collected from these mice. The specificity and the titer of the resulting sera was tested on Western blots using total insect cell extracts with or without the SERK fusion proteins.

INTRODUCTION OF THE SERK GENE INTO *PLANTA* AND THE PRODUCTION OF APOMICTIC SEED

Carrot transformation with a SERK promoter fragment/luciferase gene fusion

The binary vector pMT500 is based on the pBIN19 vector (Bevan, 1984) and contains the firefly luciferase gene downstream of a polylinker containing 5 unique restriction sites was created by uni-directional ligation of the firefly luciferase coding region followed by the polyadenylation sequence from the pea *rbcS*::E9 gene in the *Hin*dIII-*Xba*I site of the binary vector pMOG800 (kindly provided by Mogen N.V., Leiden, The Netherlands). The binary vector pMOG800 is based upon pBIN19 (Bevan, 1984) but while in pBIN19 the polylinker is flanked by the left border and the neomycin phosphotransferase (NPT II) expression cassette, the polylinker in pMOG800 is flanked by the right border and the NPT II expression cassette. From a genomic lambda clone, transcription regulating sequences from the carrot SERK gene were isolated by digestion with *Hin*dIII and DraI (SEQ ID No. 1), and cloned into the HindIII / SmaI sites of pBluescript SK+. From the resulting vector a KpnI / SstI fragment containing the SERK genomic DNA was isolated and cloned into the KpnI / SstI sites of the binary vector pMT500. The resulting DNA construct, pMT531, contained the 2200 bp genomic SERK DNA fragment as promoter sequence, the luciferase gene as vital reporter, and the E9 transcription terminator sequence.

The binary vector pMT531 was transformed by electroporation into *Agrobacterium tumefaciences* strains MOG101 and MOG301 (for transformation into carrot cells) and into *Agrobacterium tumefaciences* strain C58C1 (for transformation into *Arabidipsis thaliana* plants). Transformed colonies were selected on LB plates with 100mg/l kanamycin.

Transformation of carrot cells

The firefly luciferase coding sequence under control of the genomic carrot HindIII / DraI 2200 bp DNA fragment was introduced into carrot cells by *Agrobacterium tumefaciens* mediated transformation of hypocotyl segments. Transformation of *Daucus carota* cv. 'Amsterdamse bak' was performed by slicing one week old dark grown seedlings into segments of 10 to 20 mm. Segments were incubated for 20 minutes in a freshly prepared 10 fold diluted overnight culture of *Agrobacterium*.. The segments were dried and transferred to a modified Gamborgs B5 medium (P1 medium; S&G seeds, Enkhuizen, The Netherlands) supplemented with 2 μM 2,4-D (P1-2) and solidified with agar (Difco, Detroit, Mi, USA). After two days of culture in the dark at 25 ± 0.5 _C, segments were transferred to

solidified P1-2 medium supplemented with kanamycin (100 mg $\cdot l^{-1}$), carbenicillin (500 mg$\cdot l^{-1}$; Duchefa) and vancomycin (100 mg$\cdot l^{-1}$; Duchefa). After three weeks segments were transferred to fresh plates and transformed calli were selected after an additional three weeks. Transformed calli were grown on P1-2 plates with antibiotics for 3 weeks at a 16 hour light/8 hour darkness regime. Transformed embryogenic suspension cultures were initiated as described by transferring 0.2 g callus to 10 ml liquid P1-2 medium supplemented with 200 mg$\cdot l^{-1}$ kanamycin, 250 mg$\cdot l^{-1}$ carbenicillin and 50 mg$\cdot l^{-1}$ vancomycin. During the first weeks 1 to 3 volumes of fresh medium were added to the culture at weekly intervals. After 5 to 7 weeks cultures were subcultured to a packed cell volume of 2 ml per 50 ml medium every two weeks and incubated at a 16 hour light / 8 hour darkness regime at 25 ± 0.5 °C.

One week after transfer to kanamycin selection medium, hypocotyl segments were sprayed with luciferin to test whether luciferase expression could be detected in transformed callus shortly after transformation. A large number of hypocotyl segments showed luciferase activity at the cut edges, but did not develop calli. Instead, growth of bacteria occurred, suggesting that the luciferase activity was of bacterial origin. Six to ten weeks after transformation, calli were obtained that showed luciferase activity in variable amounts, while no bacterial growth could be observed anymore. After 12 weeks, calli measuring 5 to 10 mm in diameter were used to start suspension cultures. At this time no bacterial contamination was observed. A control transformation experiment in which luciferase expression under influence of the CaMV 35S promoter was observed in single cells and cell clusters in the suspension culture demonstrating that the luciferase protein is active in *Daucus carota* suspension cultured cells.

Cell immobilisation

One-week old high-density (10^6 - 10^7 cells$\cdot ml^{-1}$) suspension cultures were sieved through nylon sieves with successive 300, 125, 50 and 30 µm pore sizes (Monodur-PES; Verseidag Techfab, Walbeck, Germany). Single cells and cell clusters passing the last sieve are designated as < 30 µm populations. Control experiments with untransformed cells were performed with *Daucus carota* cv. 'Trophy' (S&G seeds) suspension cultures grown in P1-2 medium. Size fractionated cell populations smaller then 30 µm were immobilised in phytagel (P8196; Sigma, St Louis, Mo, USA) in petriperm dishes (Heraeus, Hanau, Germany). The bottom layer consisted of 1 ml P1-0 medium with 5 mM Ca^{2+} and 0.2 % phytagel. Two

hundred thousand cells (< 30 µm and < 50 µm populations) in B5-0 medium without Ca^{2+} supplemented with 0.1 % phytagel were poured on top of the bottom layer. For this layer B5 was applied since, at room temperature, phytagel solidified in P1 medium without Ca^{2+}. After 2 hours of solidification an additional P1-0 layer with 0.2 % phytagel was poured onto the cell layer preventing the B5 layer to move. To prevent dehydration of the phytagel layers and to supply luciferin to the cells, 0.5 ml P1-0 medium containing 0.05 µM luciferin (Promega, Madison, Wi, USA) was added after solidification. The final luciferin concentration in the culture was 0.02 µM. Luciferin detection on single cells was determined with a CCD camera for a period of 5 times one hour (Schmidt et al. (1997) Development 124: 2049-2062). After 7 days of culture, luciferin was removed from the cultures by extensive washing with P1-0 medium.

***Arabidopsis* transformation with a SERK promoter fragment/luciferase gene fusion** transformation of plants

Wildtype WS plants were grown under standard long day conditions: 16 hours light and 8 hours dark.

The first emerging influorescense was removed in order to increase the number of influorescenses. Five days later, plants were ready for vacuum infiltration.

Agrobacterium strain C58C1 containing the transformation plasmid was grown on a LB plate with 50 mg/l kanamycin, 50 mg/l rifampicin and 25 mg/l gentamycin. A single colony was used to inoculate 500 ml of LB medium containing 50 mg/l kanamycin, 50 mg/l rifampicin and 25 mg/l gentamycin. The cultures were grown O/N at 28 degrees Celcius and the resulting log phase culture (OD600 0.8) was centrifuged to pellet the cells and resuspended in 150 ml of infiltration medium (0.5x MS medium (pH 5.7) with 5% sucrose and 10 µl/l benzylaminopurine). The inflorescenses of 6 Arabidopsis plants are submerged in the infiltration suspension while he remaining parts of the plants (which are still potted) are placed upside down on meshed wire to avoid contact with the infiltration suspension.

Vacuum is applied to the whole set-up for 10 min. at 50 kPa. Plants are directly afterwards placed under standard long day conditions. After completed seed setting the seeds were surface sterilized by a 1% sodium hypochlorite soak, then thoroughly washed with sterile water and plated onto petridishes with 0.5xMS medium and 80 mg/l kanamycin in order to select for transformed seeds. After 5 days germination under long day conditions (10.000 lux), the transformed seedlings could be identified by their green color of their cotyledons

(the untransformed seedlings turn yellow), and were further grown in soil under C1 lab conditions under long day conditions. This vacuum infiltration method resulted in approximately 0.1% transformed seeds.

Transformation of a construct containing both a gene encoding kanamycin resistance and the 2200 bp (HindIII / DraI) SERK genomic DNA fused to the firefly luciferase gene into *Arabidopsis thaliana* (WS) by vacuum infiltration resulted in six different kanamycin-resistent primary transformants (I, II, III, IV, V and VI). Plants IV and VI died at the seedling stage, although they were kanamycin resistant. A T2 generation could be obtained from the four plants I, II, III and V (Figure 4). Within the siliques of the T2 generation of plants no. III and V, an early inhibition in development could be observed in appoximatetely 25-50 % of the seeds. The plants I and II did not show a reduction in the number of developing seeds.(Figure 5). Similar results were observed in a T3 generation, in which again approximately 25-50% of the seeds showed an early inhibition of normal seed development.

Arabidopsis **transformation with a AtSERK gene**

Isolation of the AtSERK genomic and cDNA clones

Using the DcSERK cDNA sequence (seq ID no. 2) as a probe, a lambda ZipLox genomic library made form Arabidopsis Landsberg erecta total genomic DNA is screened for the presence of homologous sequences. Three different lambda clones with inserts of 14, 18 and 20 kb respectively are obtained. The 14 kb clone is digested by EcoRI and the resulting fragments subcloned into pBluescript vectors. Fragments spanning the entire coding sequence of the AtSERK gene are isolated, sequenced and compared with the Daucus homologues. The resulting sequence is shown as SEQ ID NO: 20.

Using the DcSERK cDNA sequence (SEQ ID NO: 2) as a probe, a lambda ZAPII cDNA library is screened for the presence of homologous sequences. Four lambda clones are obtained and their inserts subcloned into pBluescript vectors using the helper phage excision procedure. Fragments spanning the entire AtSERK cDNA coding sequence of the AtSERK gene are isolated, sequenced and compared with the Daucus homologues. The resulting sequence is shown as SEQ ID NO: 32.

Plasmids containing promoter sequences

Arabidopsis thaliana LTP1 promoter fragment is obtained from the binary plasmid pUH1000 (Thoma, S., Hecht, U., Kipper, A., Borella, J., De Vries, S.C., Sommerville, C. (1994) Plant Physiol. 105, 35-45) by digestion with *Bam*H1 and *Hin*dIII and cloning into pBluescript SK$^-$ (pMT121).

- The CaMV 35S promoter enhanced by duplication of the -343 to -90 region (Kay *et al.*, (1987) Science 236: 1299-1302) is isolated from the pMON999 vector by digestion with *Hin*dIII and *Sst*I and cloned into the pBluescript SK$^-$ vector (pMT120).

- The promoter AtDMC1 (Klimyuk and Jones (1997) Plant Journal 11: 1-14).
Plasmid SLJ 9691 is a construct consisting of pBluescript SK+ in which the *Arabidopsis thaliana* DMC1 genomic clone (accession number U76670) is cloned into the EcoRV site. SLJ 9691 carries *Eco*RV fragments of the 5' end of the AtDMC1 gene with the following modification: a BglII site instead of the second HpaI site, two ATG codons in the first exon and an XhoI site at the ATG codon of the second exon.

- The FBP7 promoter from Petunia (Angenent et al. (1995) Plant Cell 7: 1569-1582).
The promoter of the FBP7 gene is cloned by subcloning the 0.6 kb *Hin*dIII - *X*baI genomic DNA fragment of FBP7 into the *Hin*dIII - *X*baI site of pBluescript KS-, resulting in the vector FBP201.

The pAtSERK binary vector constructs.

Based on the pBIN 19 vector, a binary vector pAtSERK is constructed for transformation of the *Arabidopsis thaliana* SERK cDNA under the control of different promoters.

The full length *Arabidopsis thaliana* cDNA clone of SERK (Seq ID No. NEW) is obtained from a pBluescript SK- plasmid. A SmaI - KpnI 2.1 kb fragment containing the AtSERK cDNA is cloned into pBIN19 SmaI - KpnI. The polyadenylation sequence from the pea rbcS::E9 gene (Millar et al., 1992), Plant Cell 4: 1075-1087) is placed downstream from the AtSERK cDNA by cloning a Klenow-filled EcoRI - HindIII E9 DNA fragment into the Klenow-filled XmaI site of the pBIN19:AtSERK vector in order to generate the binary vector pAtSERK.

Construction of plant expression vectors

The pAtSERK binary vector is used to generate the following promoter-AtSERK constructs.

- The AtLTP1 promoter is cloned in the SmaI site of the pAtSERK binary vector as a Klenow-filled *KpnI-SstI* DNA fragment to give the pAtLTP1AtSERK vector.
- The CaMV 35S promoter is cloned in the SmaI site of the pAtSERK binary vector as a Klenow-filled *KpnI-SstI* fragment to give the p35SAtSERK vector.
- The AtDMC1 promoter consisting of the BglII - XhoI 3.3kB fragment from the clone SLJ 9691 is filled in with Klenow and cloned into the SmaI site of the pAtSERK binary vector to give the pAtDMC1AtSERK vector.
- A SacI-KpnI fragment of FBP2101 is filled in with Klenow and cloned into the SmaI site of the pAtSERK binary vector to give the pFBP2101AtSERK vector.

Introduction of plant expression vectors into *Arabidopsis thaliana* plant cells

The above described vector constructs (pAtLTP1AtSERK, p35SAtSERK, pAtDMC1AtSERK, pFBP2101AtSERK) have been electrotransformed into *Agrobacterium tumifacienses* strain C58C1 as known in the art.

Wild type *Arabidopsis thaliana* WS plants are grown under standard long day conditions: 16 hours light and 8 hours dark.

The first emerging inflorescence is removed in order to increase the number of influorescences. Five days later, plants are ready for vacuum infiltration. Agrobacterium strain C58C1 containing the transformation plasmid (the pAtLTP1AtSERK vector or the p35SAtSERK or the pAtDMC1AtSERK vector or the pFBP2101AtSERK vector) is grown on a LB plate with 50 mg/l kanamycin, 50 mg/l rifampicin and 25 mg/l gentamycin. A single colony is used to inoculate 500 ml of LB medium containing 50 mg/l kanamycin, 50 mg/l rifampicin and 25 mg/l gentamycin. The cultures are grown O/N at 28 degrees Celsius and the resulting log phase culture (OD600 0.8) is centrifuged to pellet the cells and resuspended in 150 ml of infiltration medium (0.5x MS medium (pH 5.7) with 5% sucrose and 1 mg/l benzylaminopurine). The inflorescences of 6 Arabidopsis plants are submerged in the infiltration suspension while the remaining parts of the plants (which are still potted) are placed upside down on meshed wire to avoid contact with the infiltration suspension.

Vacuum is applied to the whole set-up for 10 min. at 50 kPa. Plants are directly afterwards placed under standard long day conditions. After completed seed setting the seeds are surface sterilized by a 1% sodium hypochlorite soak, then thoroughly ished with sterile water and plated onto petridishes with 0.5xMS medium and 80 mg/l kanamycin in order to select for transformed seeds. After 5 days germination under long day conditions (10.000 lux), the transformed seedlings could be identified by their green colour of their cotyledons (the untransformed seedlings turn yellow), and are further grown in soil under long day conditions. This vacuum infiltration method resulted in approximately 0.1% transformed seeds.

Expression of SERK sequences in *Arabidopsis thaliana* plant cells

The inflorescences from transgenic and not transgenic *Arabidopsis thaliana* plants are analysed by Whole mount *in situ* hybridisation analysis with AtSERK cDNA as probe. The inflorescences in different stages of development are fixed for 60 min. in PBS containing 70 mM EGTA, 4% paraformaldehyde, 0.25% glutaraldehyde, 0.1% Tween 20, and 10% DMSO. Samples are then washed, treated with proteinase K for 10 min, again washed and fixed a second time. Hybridisation solution consisted of PBS containing 0.1% Tween 20, 330 mM NaCl, 50 mg/ml heparin, and 50% deionized formamide. Hybridisation took place for 16 hours at 42°C using digoxigenin-labeled sense or antisense riboprobes (Boehringer Mannheim). After washing, the cells are treated with RNaseA and incubated with anti-digoxigenin-alkaline phosphatase conjugate (Boehringer Mannheim) which had been preabsorbed with a plant protein extract. Excess antibody is removed by washing followed by rinsing in staining buffer (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$, 1 mM levamisole) and the staining reaction is performed for 16 hours in a buffer containing NBT and BCIP. Observations are performed using a Nikon Optiphot microscope equipped with Nomarski optics.
The transformed plants show ectopic expression of SERK in the vicinity of the embryo sac.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: NOVARTIS AG
(B) STREET: Schwarzwaldallee 215
(C) CITY: Basel
(E) COUNTRY: Switzerland
(F) POSTAL CODE (ZIP): 4058
(G) TELEPHONE: +41 61 69 11 11
(H) TELEFAX: + 41 61 696 79 76
(I) TELEX: 962 991

(ii) TITLE OF INVENTION: Improvements in or relating to organic compounds

(iii) NUMBER OF SEQUENCES: 33

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6695 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Daucus carota

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3696..6617

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 3731..3802

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 3851..3979

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 4124..4211

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 4284..4357

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 4430..4528

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 4642..4757

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 4890..4967

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 5295..5803

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 6197..6339

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TCTAGATGAC GAAATCGCGC TACCTTTGAT TTNGAAATAC TAGGTTGTAG TATCTTGATT 60

AGTTTTTTGG ATATCTTGCT GTAATTTCTT TAGGAGATGC AAACGGTCTT CATTTAATAT 120

GAGCCCTTGT GACTTGACAA AAGTATCTAG CATGTTTGAT CACGAGGTAG CTAAAAAGTA 180

GCGTGTTTGA TTAAGCACAT AATATTGTAT TGGGCCTATT GGCTATCAAT GAAGTTTGAT 240

GCAAGTATAT AGCTTGTATT ATGCATGTGA TGAGGGTATA TAAAAGAAGT AAAGAACATT 300

CTCTCGTAGC ATTCATTTTT CTCTTGCCTA TAGTTAACGA GTTTTGTCAC ACATGACGTT 360

GAAACTGGAT GTGTCTGTTC TTCCATCTAA GTTTGGATTA CCTGATAGAT GCTCAACTTC 420

TTCGTCAGCC TTTTCTTTCC GATTTTTCCC AAGACAAGAT TCTTTAGTTA ATAGTTATTG 480

CTCTGGTGGC TTGTGTGCAT TTTAGGAATC TTACTCTGTT TTTTAATGGA GAAACGAAAC 540

CTACCTTTTT TTCTGTGTTC CCTTTTATGA TATCACCTGC TTGGAGGCGT TTAGACTTTA 600

TCCACCTAAA CTATTCATGT TTACCAGACA AGCTATACGT TTTATCCCCC CCCCCCGCGG 660

ACCTGNGGAC AAAAGAAGCG CTGATGAACT GATTTAATCC GTGTTTTATT ATATTACACA 720

TTGATGCTTC ATGGAGCTAA TATCTTTGGT TAAATTTCAT GTATATATAT ACCCTTCCCT 780

CTTGTGATGG CAGTGGCCCC TCGTTTAATT AGCGTACTTA ATTATCTGAT GGATACTGTA 840

TGCTTGGCAG ATGATGTCAT CAGATTATAC CATTTGTTGT GCTCTACAAA ATAAAAAACC 900

TCTATTTATG TTCATCTTTT TGGTAACAAG TAACTAATTG ATGCGCTATG TTGACAGGCG 960

ATGCATTACA CAACTTACGA ACTAGCTTGC AAGATCCCAA CAATGTCCTG CAGAGCTGGG 1020

ATCCAACCCT TGTGAACCCT TGCACATGGT TTCATGTGAC ATGTAACAAT GAAAACAGTG 1080

TTATAAGAGT GTAGGTCACT TCCCTTATTA ATTTTTTTAG CAAGTTACGA ATATTTACTC 1140

AATTGAGCAG ATGTCTCTTT AAATATTTTT CTTTAATTTC TTAGCTAAGC GGAGCATCTA 1200

TCTTAAGTAT CTCTACTGAA TTTAAGACAT AATACATTTT TTTAAAAAAT CTATTAGAGT 1260

GTTTTTTCCG CACAGCGCAC ATATATCTTT TTTCTGGTAA TTCAGACAAC CTTTCTCCCG 1320

ACGATAAAAT AATATAAGAT TAACTCCTTG AACTAATTTT TTATTTTTCT TTTCTTTTTA 1380

TGTTCTTTGC AGAAAGTTTC TTATGGTCTT TTGTGAAAAG TACATTCTAT GATAATTTTT 1440

TGGCAACTCA TATAAATTTA TATATATTCC ATGTAGTTAT AAGTTAAAAA AAGCTTCCTA 1500

TTAATTCCAA GATAGAGGTT CATTTTTATA GTTTGGGCAT CCATGAGTTT TTGAAAATGT 1560

CAGAAATTTT GTTGAGTTAA TTTTACTTAC CAACTTTTAT GGCGTCATGC AGTGATCTTG 1620

GGAATGCAGC ATTATCTGGT CAATTGGTTC CTCTTGGCCA GTTGAAAAAT TTACAATACT 1680

TGTAAGACCA TATCACTTGG AATGCTTTAG TTTTTATACA GCACAATGCT TTCAATATCT 1740

GTTAAAAGTG TGAAAAAGTT GACTTTCTAG CTTCAGCAGT TGTTCGGATA ATATCTATGA 1800

AGCACTTAAA AGGCTGGGCA ATTTTTTTGT TATTATTTCA AATATTGTTA ATTGTTACTA 1860

CTTAATATGA TAAACTGATT TAACTCCTCA TGATTGGTCT CAGTCCAATG TGCCCTCATT 1920

AGTCACATNA TAAAATTGGN GGGTTGGACA AATATAACTT CTTTTCTTAA GGTCCAGAAA 1980

GAGCACTTAT CAACCTTGTC TAGCGCATAA CGTCACAGTG GGTCAGTCAC GGGCTATCCA 2040

GTTTGGGGAG GTTTTAATGA GCACTTATTT ACCTTGTCTT TTAAACGTCT GAGGATGTTA 2100

TTAAAGTCTG CATCATTCAG AGTTTAAATT AGCACTTTCA GTTGTATTAT GAATGGTACA 2160

TGAAAGATAC ATATCTTAAT GTTCCTATGC CTGTTTCAAC ATGTCTCTAA TATTCTGTTA 2220

TCTTTGTCAT CTTAAAAATG GCACTGATTA AAATGTGAGA AAGGTAGTCT TCCAATACCA 2280

TTTCATGTAT ACCAGAGAAT ATCATAATTT TTTTAAATCA TAAGTTGGGC CCTAGAGTTT 2340

TCTCAGTATT GGTCTATTTA TATTTTCCAC CATTTAGAAC TGTGTTGTCA GATGAAAATC 2400

TTGGACTTCC ACAGAAGATC TTATAGTAAA AGTATTCTTT AGATCTGATG ATGAAAGTTG 2460

TCATGGTGTG GCCTGTCCCA GAATTTAAAT CAATCCCATG TCACATGTTT GTTGATCTGA 2520

CTACTCACTG TTAATCGAAG AGTAACTATT TGTGAATTAA ATGCTTTTTT TTTTGTTCTT 2580

CATGCTTAGC GTTATAAAGG TCTACGTCTG ACTATGGTTT TTAACATGTT ATAGTTTTGT 2640

ACTGACAAGT TTAAAGTTTC TCTTGTTTAC GAATTAAGAA TATATAATAT AAAACGCTTT 2700

AACTTTCTCT GTGGAAGGTG TTCTTACCTT TTTATATATA TATATAGATA CTCAGACTCT 2760

GCTGGCAATT ATATCTTACG AACTTACGAG TATACAGAAC TTGTATATTA GGTTCAGATG 2820

AGTGGCTGTA GTAGAACACC TTAAGCAAGA ACTTAATCAT GAGGTTTCAA CCTTTTAACT 2880

TTCTTTTTAG ATTTTTTCAA GTTTATGGAA AATTGTACCT CATGATCGTG GTTTCTTTCC 2940

ATAAACTTTC CATATAAGTC CGTTTCTTGA CGTTTTCATG TAAGCTGTTG ACGAGTGATT 3000

ATTAGCGGTT CTTTCAATAA TCATAATGTG TCTCACTTTG ATGAGGCCTG TACTTATTAT 3060

TGCACCTTGC ACTTAACCTT GATCCTCATG TCATCTTGAT TGTCATAGTC TACTAACCGA 3120

GTTGAACATG GTTTATCATG TCTTTTGAGG TAACAATGTA GCTTTCACCT CTGTCCTTGA 3180

TATAGGTTTA AGGCTTGCAC CTCCCACTAG CCTTTCGTTG TTTTATTCAC AGTTCACACA 3240

CCTACTAGCA CTGTTCACCT CTAGTCTTTT GTCCGCAAAT AGTAAGAAGT TTCTTTCGCA 3300

TAATAGTGGA TGATCATTTA AGAAATAGTG AATCAAATTA TCGTGTTATT GTGTTTGTAC 3360

TTTGGAATTA AATGAGTTGC TGAACATTGT TGCTGTTTAT CGTTGTCAAG GCTTTGCCAA 3420

GGAAGGCGAT TAGTAAGAGT GGGCATCCAA GCGCCTTTAT CTTGAAGGGG CGGGCGGCAC 3480

GTTGTGGATT CTGGGTGTCT ATTAGAGGAC ATTATCTATA TATACTGATT ATTTATTAGA 3540

ATATAAATCA ACTACTATAT TTTTCTTTGT AATGTTTATA TAGAAATCCC ACTCGTAAAC 3600

TTGACAAATA CCATTGAAAT ATTTGAACCT AATTAATTAG TAGTGTCAGG TTTAAATTCA 3660

AACTCATTTA ATTTTACTTT AAAAAATAAT TCTATATGAA TCGTAACAGT ATAAATATAT 3720

TAAATTACAT GTATGTGTGC CTATATATAG CTGAATGTCT AATAGACTCC AAGACGGCTG 3780

CTCTTACTGC CTAGGCGTCC AGGCAGTTCA CTGATGCTTA CCTTGACAAA TATGGGGTTC 3840

GTATGACATT GTTGGGGATC CCTATCACTG GATTCCTGTT TTGCTGACCC TCTGTTCAAT 3900

TGATTTTCAT TGATGTAGTA TTACTAGTTT TATAAATATT CTTTATTGCA ATAATTTAAC 3960

TGGAGTTTAA CAATGACAGG GAGCTTTACA GCAATAACAT AAGTGGACCA ATTCCTAGTG 4020

ATCTTGGGAA TCTGACAAAT TTGGTGAGCT TGGACCTATA CATGAATAGC TTCTCTGGAC 4080

CTATACCGGA CACATTAGGA AAGCTTACAA GGCTAAGATT CTTGTATGAC TACAAATCTT 4140

CACTAGTTTT TAACTTAATG CAATTTGATT ATCCTTTCAA GTGATTGATT ATATCACAAA 4200

TTACTGGATA GGCGTCTCAA CAACAACTGC CTCTCTGGTC CAATTCCAAT GTCACTGACT 4260

AATATTACAA CTCTTCAAGT CCTGTAAGTA TTCCGACCTT TCCAGATAGT TTTGTTGTTG 4320

TGGATGTTTC AATTTTAATA CTAAATATGT TCATCAGGGA TTTATCAAAC AATCGGCTAT 4380

CAGGACCAGT ACCGGATAAT GGCTCATTTT CTTTGTTTAC ACCTATCAGG TTTAATGCTA 4440

GTAATATCTT TAATATTATG GTTCTTACTT CTACTGCGAA AGCTATGATA ATATTTTTTT 4500

TCTCCTTCAT ATATTATCAC TTTCGCAGTT TTGGCAATAA TTTGAATTTA TGTGGACCTG 4560

TAACTGGGAG GCCCTGCCCT GGATCTCCCC CATTTTCTCC ACCACCTCCG TTCATCCCAC 4620

CATCAACAGT ACAGCCTCCA GGTGATTTAG TTTTTATATT AATTCCCGTA ATTAATTTTA 4680

TGACTGTAAA AATTGGTGTT AATTTCACCA GTTGCGAATA AAGTATTTTC CTTCTTTCTC 4740

TTCTTATTAT TATGAAGGAC AAAATGGTCC CACTGGAGCT ATTGCTGGGG GAGTAGCTGC 4800

TGGTGCTGCT TTACTGTTTG CTGCACCTGC AATGGCATTT GCATGGTGGC GGAGAAGAAA 4860

ACCGCGAGAA CATTTCTTTG ATGTGCCAGG TTAGTCCTGT AAATAGATAT CTATTGAAGC 4920

GCTTACTGTC TGTGGACTTT GTTTTCACTG TCATTAGTTA ACTTCAGCTG AAGAGGACCC 4980

AGAAGTGCAC CTTGGTCAAC TGAAGAGGTT TTCTCTGCGA GAATTGCAAG TCGCAACGGA 5040

TACTTTTAGT ACCATCCTTG GAAGAGGTGG ATTTGGTAAG GTGTATAAGG GACGCCTTGC 5100

TGATGGCTCA CTTGTAGCAG TTAAAAGGCT TAAAGAAGAA CGAACACCAG GTGGCGAGCT 5160

GCAGTTTCAA ACAGAAGTGG AAATGATTAG CATGGCTGTG CATCGAAATC TTCTGCGTCT 5220

ACGTGGTTTC TGCATGACAC CTACCGAGCG GCTTCTTGTA TATCCATACA TGGCTAATGG 5280

AAGTGTTGCG TCATGTTTAA GAGGTATCTC AGTTACAATT ACCATAACTT GCCAGAAGTT 5340

TGTTTGATTA AAAATGAAAT ATAACTCCCT ACACTATGTT AAGGTGTTAT AATTTCTGAG 5400

CAGATCTTAT TTCCCATTGC AAGATACCAG TTATTATTGT TTTTTCTGTA ATTGATACCG 5460

GTTATATTTC TTTCTTGTAT TTGGTTATAT GCAAGGATTT CGAGTCTAAT AAGTTATCAA 5520

ACTGGATGCT ATGTTTATTC TGCAATTGAA TTCTTGCTTC ATGTGCCAAA ATATATATGA 5580

TTCAACTTGG AATCATCTTA TAATATACTG TGTAAAGTCA GCTGTTGACT TTCATCATTA 5640

ATTAGTCTTC ATAAATCAGA ATCTGCCTAG TGAGCTTTAC CGACATACTC TAAACCTTTC 5700

TTATGGCCCT GTATATAATC GTCCCACTTA CTTTATTCAG TTTGTCTGCT CTCTGAATTT 5760

TTGATCTGTA CATTGTGATG TCTTGTTTTC ATCAAATGTA GAGCGTCAGC CATCAGAACC 5820

TCCCCCTGAT TGGCCAACTA GGGAGAGGAT TGCACTAGGA TCTTCTAGGG GCCTATCTAA 5880

ATTGCATGAC CATTGTGATC CCAAGATTAT CCATCGCGAT GTAAAAGCTG CAAATATATT 5940

ATTGGACGAA GAATTTGAGG CTGTTGTAGG TGATTTTGGG TTAGCTAGGC TCATGGATTA 6000

CAAGGATACC CATGTTACGA CTGCTGTAAG GGGTACCATT GGGCACATAG CTCCCGAGTA 6060

CCTCTCGACT GGAAAGTCAT CAGAGAAGAC CGATGTCTTT GGTTATGGGA TAATGCTCCT 6120

AGAGCTCATT ACTGGACAGA GGGCTTTTGA TCTTGCTCGC CTTGCGAACG ATGATGATGT 6180

TATGTTGTTG GATTGGGTAT GTGTCCCGGG TGTTCCTTTG GTTAATTATT TCACATATTA 6240

GTGCTTACTA CTTTGTTGTG GCCCTTTGTT TTTATTTCCT GCCTGTATTT GATTCTTAGT 6300

CATGTTATGC ATATTGACCT GCTTTGCAAT GTCTTTTAGG TTAAAAGCCT TTTGAAAGAG 6360

AAAAAGTTGG AGATGCTGGT CGATCCTGAC CTGCAGAACA ATTACATTGA CACAGAAGTT 6420

GAGCAGCTTA TTCAAGTAGC ATTACTCTGT ACCCAGGGTT CGCCAATGGA GCGGCCTAAG 6480

ATGTCAGAGG TAGTCCGAAT GCTTGAAGGT GATGGCCTTG CAGAAAAGTG GGACGAGTGG 6540

CAAAAAGTTG AAGTCATCCA TCAAGACGTA GAATTAGCTC CACATCGAAC TTCTGAATGG 6600

ATCCTAGACT CGACAGATAA CTTGCATGCT TTTGAATTAT CTGGTCCAAG ATAAACAGCA 6660

TATAAAATGT AATGAAATTA ATATTTTTTA TGGTT 6695

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1815 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Daucus carota

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 94..1752

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
GACAAATACC ATTGAAATAT TTGAACCTAA TTAATTAGTA GTGTCAGGTT TAAATTCAAA        60

CTCATTTAAT TTTACTTTAA AAAATAATTC TAT ATG AAT CGT AAC AGT ATA AAT        114
                                     Met Asn Arg Asn Ser Ile Asn
                                       1               5

ATA TTA AAT TAC ATG CAG TTC ACT GAT GCT TAC CTT GAC AAA TAT GGG        162
Ile Leu Asn Tyr Met Gln Phe Thr Asp Ala Tyr Leu Asp Lys Tyr Gly
        10                  15                  20

GTT CTT ATG ACA TTG GAG CTT TAC AGC AAT AAC ATA AGT GGA CCA ATT        210
Val Leu Met Thr Leu Glu Leu Tyr Ser Asn Asn Ile Ser Gly Pro Ile
    25                  30                  35

CCT AGT GAT CTT GGG AAT CTG ACA AAT TTG GTG AGC TTG GAC CTA TAC        258
Pro Ser Asp Leu Gly Asn Leu Thr Asn Leu Val Ser Leu Asp Leu Tyr
 40                  45                  50                  55

ATG AAT AGC TTC TCT GGA CCT ATA CCG GAC ACA TTA GGA AAG CTT ACA        306
Met Asn Ser Phe Ser Gly Pro Ile Pro Asp Thr Leu Gly Lys Leu Thr
                60                  65                  70

AGG CTA AGA TTC TTG CGT CTC AAC AAC AAC AGC CTC TCT GGT CCA ATT        354
Arg Leu Arg Phe Leu Arg Leu Asn Asn Asn Ser Leu Ser Gly Pro Ile
            75                  80                  85

CCA ATG TCA CTG ACT AAT ATT ACA ACT CTT CAA GTC CTG GAT TTA TCA        402
Pro Met Ser Leu Thr Asn Ile Thr Thr Leu Gln Val Leu Asp Leu Ser
        90                  95                 100

AAC AAT CGG CTA TCA GGA CCA GTA CCG GAT AAT GGC TCA TTT TCT TTG        450
Asn Asn Arg Leu Ser Gly Pro Val Pro Asp Asn Gly Ser Phe Ser Leu
    105                 110                 115

TTT ACA CCT ATC AGT TTT GCC AAT AAT TTG AAT TTA TGT GGA CCC GTA        498
```

Phe Thr Pro Ile Ser Phe Ala Asn Asn Leu Asn Leu Cys Gly Pro Val
120 125 130 135

ACT GGG AGG CCC TGC CCT GGA TCT CCC CCA TTT TCG CCA CCA CCT CCG 546
Thr Gly Arg Pro Cys Pro Gly Ser Pro Pro Phe Ser Pro Pro Pro Pro
140 145 150

TTC ATC CCA CCA TCA ACA GTA CAG CCT CCA GGA CAA AAT GGT CCC ACT 594
Phe Ile Pro Pro Ser Thr Val Gln Pro Pro Gly Gln Asn Gly Pro Thr
155 160 165

GGA GCT ATT GCT GGG GGA GTA GCT GCT GGT GCT GCT TTA CTG TTT GCT 642
Gly Ala Ile Ala Gly Gly Val Ala Ala Gly Ala Ala Leu Leu Phe Ala
170 175 180

GCA CCT GCA ATG GCA TTT GCA TGG TGG CGG AGA AGA AAA CCG CGA GAA 690
Ala Pro Ala Met Ala Phe Ala Trp Trp Arg Arg Arg Lys Pro Arg Glu
185 190 195

CAT TTC TTT GAT GTG CCA GCT GAA GAG GAC CCA GAA GTG CAC CTT GGT 738
His Phe Phe Asp Val Pro Ala Glu Glu Asp Pro Glu Val His Leu Gly
200 205 210 215

CAA CTG AAG AGG TTT TCT CTG CGA GAA TTG CAA GTC GCA ACG GAT ACT 786
Gln Leu Lys Arg Phe Ser Leu Arg Glu Leu Gln Val Ala Thr Asp Thr
220 225 230

TTT AGT ACC ATA CTT GGA AGA GGT GGA TTT GGT AAG GTG TAT AAG GGA 834
Phe Ser Thr Ile Leu Gly Arg Gly Gly Phe Gly Lys Val Tyr Lys Gly
235 240 245

CGC CTT GCT GAT GGC TCA CTT GTA GCA GTT AAA AGG CTT AAA GAA GAA 882
Arg Leu Ala Asp Gly Ser Leu Val Ala Val Lys Arg Leu Lys Glu Glu
250 255 260

CGA ACA CCA GGT GGT GAG CTG CAG TTT CAA ACA GAG GTG GAA ATG ATT 930
Arg Thr Pro Gly Gly Glu Leu Gln Phe Gln Thr Glu Val Glu Met Ile

265 270 275

AGC ATG GCT GTG CAT CGA AAT CTT CTG CGT CTA CGT GGT TTC TGC ATG 978
Ser Met Ala Val His Arg Asn Leu Leu Arg Leu Arg Gly Phe Cys Met
280 285 290 295

ACA CCA ACA GAG CGG CTT CTT GTA TAT CCA TAC ATG GCT AAT GGA AGT 1026
Thr Pro Thr Glu Arg Leu Leu Val Tyr Pro Tyr Met Ala Asn Gly Ser
300 305 310

GTT GCG TCG TGT TTA AGA GAG CGT CAG CCA TCA GAA CCT CCC CTT GAT 1074
Val Ala Ser Cys Leu Arg Glu Arg Gln Pro Ser Glu Pro Pro Leu Asp
315 320 325

TGG CCA ACT AGG AAG AGG ATT GCA CTA GGA TCT GCT AGG GGG CTT TCT 1122
Trp Pro Thr Arg Lys Arg Ile Ala Leu Gly Ser Ala Arg Gly Leu Ser
330 335 340

TAT TTG CAT GAC CAT TGT GAT CCC AAG ATT ATC CAT CGT GAT GTA AAA 1170
Tyr Leu His Asp His Cys Asp Pro Lys Ile Ile His Arg Asp Val Lys
345 350 355

GCT GCA AAT ATA TTA TTG GAC GAA GAA TTT GAG GCT GTT GTA GGT GAT 1218
Ala Ala Asn Ile Leu Leu Asp Glu Glu Phe Glu Ala Val Val Gly Asp
360 365 370 375

TTT GGG TTA GCT AGG CTC ATG GAT TAC AAG GAT ACC CAT GTT ACA ACT 1266
Phe Gly Leu Ala Arg Leu Met Asp Tyr Lys Asp Thr His Val Thr Thr
380 385 390

GCT GTA AGG GGT ACC TTG GGC TAC ATA GCT CCC GAG TAC CTC TCG ACT 1314
Ala Val Arg Gly Thr Leu Gly Tyr Ile Ala Pro Glu Tyr Leu Ser Thr
395 400 405

GGA AAG TCA TCA GAG AAG ACC GAT GTC TTT GGT TAT GGG ATT ATG CTC 1362
Gly Lys Ser Ser Glu Lys Thr Asp Val Phe Gly Tyr Gly Ile Met Leu
410 415 420

TTA GAG CTC ATT ACT GGA CAG AGA GCT TTT GAT CTT GCT CGC CTT GCG 1410
Leu Glu Leu Ile Thr Gly Gln Arg Ala Phe Asp Leu Ala Arg Leu Ala
425 430 435

AAC GAT GAT GAT GTT ATG TTG TTG GAT TGG GTT AAA AGC CTT TTG AAA 1458
Asn Asp Asp Asp Val Met Leu Leu Asp Trp Val Lys Ser Leu Leu Lys
440 445 450 455

GAG AAA AAG TTG GAG ATG CTG GTC GAT CCT GAC CTG GAG AAC AAT TAC 1506
Glu Lys Lys Leu Glu Met Leu Val Asp Pro Asp Leu Glu Asn Asn Tyr
460 465 470

ATT GAC ACA GAA GTT GAG CAG CTT ATT CAA GTA GCA TTA CTC TGT ACC 1554
Ile Asp Thr Glu Val Glu Gln Leu Ile Gln Val Ala Leu Leu Cys Thr
475 480 485

CAG GGT TCG CCA ATG GAG CGG CCT AAG ATG TCA GAG GTA GTC CGA ATG 1602
Gln Gly Ser Pro Met Glu Arg Pro Lys Met Ser Glu Val Val Arg Met
490 495 500

CTT GAA GGT GAT GGC CTT GCA GAA AAG TGG GAC GAG TGG CAA AAA GTA 1650
Leu Glu Gly Asp Gly Leu Ala Glu Lys Trp Asp Glu Trp Gln Lys Val
505 510 515

GAA GTC ATC CAT CAA GAC GTA GAA TTA GCT CCA CAT CGA ACT TCT GAA 1698
Glu Val Ile His Gln Asp Val Glu Leu Ala Pro His Arg Thr Ser Glu
520 525 530 535

TGG ATC CTA GAC TCG ACA GAT AAC TTG CAT GCT TTT GAA TTA TCT GGT 1746
Trp Ile Leu Asp Ser Thr Asp Asn Leu His Ala Phe Glu Leu Ser Gly
540 545 550

CCA AGA TAAACAGCAT ATAAAATGTG AATGAAATTA ATATTTTTTA TGGTTAAAAA 1802
Pro Arg

AAAAAAAAAA AAA 1815

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 553 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
Met Asn Arg Asn Ser Ile Asn Ile Leu Asn Tyr Met Gln Phe Thr Asp
  1               5                  10                  15

Ala Tyr Leu Asp Lys Tyr Gly Val Leu Met Thr Leu Glu Leu Tyr Ser
             20                  25                  30

Asn Asn Ile Ser Gly Pro Ile Pro Ser Asp Leu Gly Asn Leu Thr Asn
         35                  40                  45

Leu Val Ser Leu Asp Leu Tyr Met Asn Ser Phe Ser Gly Pro Ile Pro
     50                  55                  60

Asp Thr Leu Gly Lys Leu Thr Arg Leu Arg Phe Leu Arg Leu Asn Asn
 65                  70                  75                  80

Asn Ser Leu Ser Gly Pro Ile Pro Met Ser Leu Thr Asn Ile Thr Thr
                 85                  90                  95

Leu Gln Val Leu Asp Leu Ser Asn Asn Arg Leu Ser Gly Pro Val Pro
            100                 105                 110

Asp Asn Gly Ser Phe Ser Leu Phe Thr Pro Ile Ser Phe Ala Asn Asn
        115                 120                 125
```

Leu Asn Leu Cys Gly Pro Val Thr Gly Arg Pro Cys Pro Gly Ser Pro
130 135 140

Pro Phe Ser Pro Pro Pro Pro Phe Ile Pro Pro Ser Thr Val Gln Pro
145 150 155 160

Pro Gly Gln Asn Gly Pro Thr Gly Ala Ile Ala Gly Gly Val Ala Ala
165 170 175

Gly Ala Ala Leu Leu Phe Ala Ala Pro Ala Met Ala Phe Ala Trp Trp
180 185 190

Arg Arg Arg Lys Pro Arg Glu His Phe Phe Asp Val Pro Ala Glu Glu
195 200 205

Asp Pro Glu Val His Leu Gly Gln Leu Lys Arg Phe Ser Leu Arg Glu
210 215 220

Leu Gln Val Ala Thr Asp Thr Phe Ser Thr Ile Leu Gly Arg Gly Gly
225 230 235 240

Phe Gly Lys Val Tyr Lys Gly Arg Leu Ala Asp Gly Ser Leu Val Ala
245 250 255

Val Lys Arg Leu Lys Glu Glu Arg Thr Pro Gly Gly Glu Leu Gln Phe
260 265 270

Gln Thr Glu Val Glu Met Ile Ser Met Ala Val His Arg Asn Leu Leu
275 280 285

Arg Leu Arg Gly Phe Cys Met Thr Pro Thr Glu Arg Leu Leu Val Tyr
290 295 300

Pro Tyr Met Ala Asn Gly Ser Val Ala Ser Cys Leu Arg Glu Arg Gln
305 310 315 320

Pro Ser Glu Pro Pro Leu Asp Trp Pro Thr Arg Lys Arg Ile Ala Leu

325 330 335

Gly Ser Ala Arg Gly Leu Ser Tyr Leu His Asp His Cys Asp Pro Lys
340 345 350

Ile Ile His Arg Asp Val Lys Ala Ala Asn Ile Leu Leu Asp Glu Glu
355 360 365

Phe Glu Ala Val Val Gly Asp Phe Gly Leu Ala Arg Leu Met Asp Tyr
370 375 380

Lys Asp Thr His Val Thr Thr Ala Val Arg Gly Thr Leu Gly Tyr Ile
385 390 395 400

Ala Pro Glu Tyr Leu Ser Thr Gly Lys Ser Ser Glu Lys Thr Asp Val
405 410 415

Phe Gly Tyr Gly Ile Met Leu Leu Glu Leu Ile Thr Gly Gln Arg Ala
420 425 430

Phe Asp Leu Ala Arg Leu Ala Asn Asp Asp Asp Val Met Leu Leu Asp
435 440 445

Trp Val Lys Ser Leu Leu Lys Glu Lys Lys Leu Glu Met Leu Val Asp
450 455 460

Pro Asp Leu Glu Asn Asn Tyr Ile Asp Thr Glu Val Glu Gln Leu Ile
465 470 475 480

Gln Val Ala Leu Leu Cys Thr Gln Gly Ser Pro Met Glu Arg Pro Lys
485 490 495

Met Ser Glu Val Val Arg Met Leu Glu Gly Asp Gly Leu Ala Glu Lys
500 505 510

Trp Asp Glu Trp Gln Lys Val Glu Val Ile His Gln Asp Val Glu Leu
515 520 525

Ala Pro His Arg Thr Ser Glu Trp Ile Leu Asp Ser Thr Asp Asn Leu
530 535 540

His Ala Phe Glu Leu Ser Gly Pro Arg
545 550

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TTTTTTTTTT TGC 13

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGGATCTAAG 10

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

ACACGTGGTC 10

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TCAGCACAGG 10

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TTTTTTTTTT TCTG 14

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TTTTTTTTTT TCA 13

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GACATCGTCC 10

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CCCTACTGGT 10

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

ACACGTGGTC 10

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GGTGACTGTC 10

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TCTTGGACCA GATAATTC 18

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CTCTGATGAC TTTCCAGTC 19

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

AATGGCATTT GCATGG 16

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Daucus carota

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Ser Pro Pro Pro Pro
1 5

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Daucus carota

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

His Arg Asp Val Lys Ala Ala Asn

1 5

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Daucus carota

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Gly Thr Leu Gly Tyr Ile Ala Pro Glu
1 5

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4081 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(vii) IMMEDIATE SOURCE:

(B) CLONE: Arabidopsis SERK gene

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 1280..1367

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 1796..1928

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 2014..2085

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 2203..2346

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 2450..2521

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 2617..2688

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 2772..2884

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 3015..3146

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 3305..3646

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 3760..4081

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

```
TCTAGAAACC TTTTGATCAT AATGAAAATA AAGAGTCCAT CCACCACATG GGGTAAGCAT      60
AATGTGTGAT ATTTAAAGGG TAACAAATGT AATCTGCTTT TTATTTTACT TTTTACCTCT     120
ACTCAAATTG TATGGGCAGT TTTTTTTTTT TTTTAAATGA TAAGACAAGT ATCTGTTTAA     180
TGGTATTGTG ATGAAACAGT AGTAAAGTCA TATCGGGCAC GCCATACTAC TTCCACAGTG     240
GAACTTGGCC AAATTTTGTC TTTGCCGTCT CTACAGTTTC TTCCACCAAA TTTTTTGTTG     300
ACAAAACTCA AATCTTTCAA TCTCATCTCT GCCAAAGTTG GGTTTAGAAA GAATATCAGC     360
AAACACTAAT ATCTTTATTG TTGCATGGTT TATCAATCAC AAAATTCACA ACCATTGTAA     420
AAAAAAATTC ACATTTTTGG TATGAGATTG CTCACATGAT AGTGAACCTC TTTAACATTT     480
TAACTTTACT TTCATAAATA CGGGATTACG AATCTTACTT GCATTAAAAA TTTAGAAAAG     540
GTTTTTCTAC TTAAAGAAAA AAGGGACCCA ACAGAGAGAG GTTTGACCAG GAGAAACGGG     600
TGCATAGCCT TAAGAGCTTT CAACTACTTT ACCCCAAACC CAAAGCGATG TCACTTTCAA     660
CCATCTCTTC TCTCCCCCGA ACCCGTTTTT TTGACCGGTC AGTTCGGGCA GCAGCACCGT     720
TACGGGCAGC TTATATTCCT CGTCTTCCTC CTCTACACCA CTGCATGCCC ATAAATAAAG     780
```

CCCGTTGAGA TCTTTAAAAA TATTAAATAA TATATCAACG AAAAAGCTAT TTTATTCATA 840

AGAAGAAAAA GAGAGGAACA ACAACAACAC ACTAATCATA GTTTCTCTGG CAGGCTTGTT 900

GTTGCGGCTT AATAAAAAGC TCTTTTGTTA TTATTACTTC ACGTAGATTT TCCCCAAAAA 960

GCTCTTATTT TTTTGTTTAA AAAAAAAAGT TTCATCTTTA TTCAACTTTT GTTTTACAGT 1020

GTGTGTGTGA GAGAGAGAGT GTGGTTTGAT TGAGGAAAGA CGACGACGAG AACGCCGGAG 1080

AATTAGGATT TTTATTTTAT TTTTTACTCT TTGTTTGTTT TAATGCTAAT GGGTTTTTAA 1140

AAGGGTTATC GAAAAAATGA GTGAGTTTGT GTTGAGGTTG TCTCTGTAAA GTGTTAATGG 1200

TGGTGATTTT CGGAAGTTAG GGTTTTCTCG GATCTGAAGA GATCAAATCA AGATTCGAAA 1260

TTTAGCATTG TTGTTTGAAA TGGAGTCGAG TTATGTGGTG TTTATCTTAC TTTCACTGAT 1320

CTTACTTCCG AATCATTCAC TGTGGCTTGC TTCTGCTAAT TTGGAAGGTT CGTGGTTACT 1380

CAATTACTCA GCTTTACTCG TTTCTCAATT ACTTTCTCGA TTCTTTTTTA TTTGGAGGTG 1440

AATCGCTATC TTTAGTGTCT GCATTTTGAT TTATGAAAAT TGTTGTTGTT CTTTGTATTT 1500

GTAAGATTTA GTGGCTAGTA CTTTGAATAC ACTGTTTTGC TTTTCTTGTT CAGATCAACT 1560

TTGTATATTG TAAAGGCATG TTCTTTGGGT TGAAAAGCTG GGTTATTTGA TATCTTAAGA 1620

TTGATGTTGT TGATCCAAAC ATTCTCTGAA AGACTTCATT TGTTTTTGGT TTTGTAAAGA 1680

ATTTGTTTAA TTATTAGCCT CTAATCTCAG AGAGGCCTGT TTGAATAGTT CTCTCTTGAA 1740

ATTAGACTTT TCACCAATTG ATGCTAATTG TGTAGATTTG TTGTTCTTGT TATAGGTGAT 1800

GCTTTGCATA CTTTGAGGGT TACTCTAGTT GATCCAAACA ATGTCTTGCA GAGCTGGGAT 1860

CCTACGCTAG TGAATCCTTG CACATGGTTC CATGTCACTT GCAACAACGA GAACAGTGTC 1920

ATAAGAGTGT AAAGCTTTCT TCTACTAATC CCACTTTTTA AACTTTGACC TCAGCGTGGT 1980

TACCGACATT TTTGTTTCTT TTGTCAAATA CAGTGATTTG GGGAATGCAG AGTTATCTGG 2040

CCATTTAGTT CCAGAGCTTG GTGTGCTCAA GAATTTGCAG TATTTGTAAG TTCCACTTAT 2100

GCATCATGCT TTAACAAAAC AAATCCAAGA TTTGACAGAA GAAGCACTGG AGTTACCTTT 2160

TGTAATTGAA ATCTTTTTAA CAAGTTTCTT ATTTTCTTAC AGGGAGCTTT ACAGTAACAA 2220

CATAACTGGC CCGATTCCTA GTAATCTTGG AAATCTGACA AACTTAGTGA GTTTGGATCT 2280

TTACTTAAAC AGCTTCTCCG GTCCTATTCC GGAATCATTG GGAAAGCTTT CAAAGCTGAG 2340

ATTTCTGTGA GTATACATAT GCTTTACCGG CTCAGTTACA GTCTTTGTTT AATCTTAGGT 2400

TTTGTTCCAA TTTTTGACTC TTTGCTGAAA ATTTTACATG CAAGAATAGC CGGCTTAACA 2460

ACAACAGTCT CACTGGGTCA ATTCCTATGT CACTGACCAA TATTACTACC CTTCAAGTGT 2520

TGTGAGTCCT CTCATTAACT TTCATTTATG TCTACTTCAT TCTCCCTCAG TTGATTTGTT 2580

GAGTTAATGC ACTTAACCTT GATGGATGCA ACACAGAGAT CTATCAAATA ACAGACTCTC 2640

TGGTTCAGTT CCTGACAATG GCTCCTTCTC ACTCTTCACA CCCATCAGGT TCTATGATTT 2700

ATCCTCTTCA GTTATTTCAG TTGTTGTGTC AGTGTCTGAA CTTATTCTGA AACTTTCATT 2760

TCCTTGTGCA GTTTTGCTAA TAACTTAGAC CTATGTGGAC CTGTTACAAG TCACCCATGT 2820

CCTGGATCTC CCCCGTTTTC TCCTCCACCA CCTTTTATTC AACCTCCCCC AGTTTCCACC 2880

CCGAGTAAGC CTCCTCTTTT TAGTTTACAT TATAGGAAAC AGAAGATGAA ATCTTTGCTT 2940

CTCTGTCAAT CCTTTTTCTC ATATAACTCA TCTTGCCAAT AAGGCAATAA CCAAATGATC 3000

TAATTTGATT TCAGGTGGGT ATGGTATAAC TGGAGCAATA GCTGGTGGAG TTGCTGCAGG 3060

TGCTGCTTTG CTCTTTGCTG CTCCTGCAAT AGCCTTTGCT TGGTGGCGAC GAAGAAAGCC 3120

ACTAGATATT TTCTTCGATG TGCCTGGTGA GTTTATTATT CGCATTAGTT TCTGTTCTTA 3180

GCCAGCAATT TTGTTTTGCA GAAAAGTATT GGAACAACTG TTAATGAAAA TCAATACATA 3240

AGTCATTGTT TTTTAAGTTA CAAACTCTTT TGAGTAAAAT CTCGATTGCA AAATCTCTAT 3300

GCAGCCGAAG AAGATCCAGA AGTTCATCTG GGACAGCTCA AGAGGTTTTC TTTGCGGGAG 3360

CTACAAGTGG CGAGTGATGG GTTTAGTAAC AAGAACATTT TGGGCAGAGG TGGGTTTGGG 3420

AAAGTCTACA AGGGACGCTT GGCAGACGGA ACTCTTGTTG CTGTCAAGAG ACTGAAGGAA 3480

GAGCGAACTC CAGGTGGAGA GCTCCAGTTT CAAACAGAAG TAGAGATGAT AAGTATGGCA 3540

GTTCATCGAA ACCTGTTGAG ATTACGAGGT TTCTGTATGA CACCGACCGA GAGATTGCTT 3600

GTGTATCCTT ACATGGCCAA TGGAAGTGTT GCTTCGTGTC TCAGAGGTAA AAACTAAACA 3660

ATTAAACATC TTGTGCTCTC TCTCAATTAC TTTGACGTGA AGTGTTTTTT CATGTTTTCC 3720

TTTATGGGTT CATAATTGTT GGTTACACTA ATGACACAGA GAGGCCACCG TCACAACCTC 3780

CGCTTGATTG GCCAACGCGG AAGAGAATCG CGCTAGGCTC AGCTCGAGGT TTGTCTTACC 3840

TACATGATCA CTGCGATCCG AAGATCATTC ACCGTGACGT AAAAGCAGCA AACATCCTCT 3900

TAGACGAAGA ATTCGAAGCG GTTGTTGGAG ATTTCGGGTT GGCAAAGCTA ATGGACTATA 3960

AAGACACTCA CGTGACAACA GCAGTCCGTG GCACCATCGG TCACATCGCT CCAGAATATC 4020

TCTCAACCGG AAAATCTTCA GAGAAAACCG ACGTTTTCGG ATACGGAATC ATGCTTCTAG 4080

A 4081

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 494 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

```
Met Glu Ser Ser Tyr Val Val Phe Ile Leu Leu Ser Leu Ile Leu Leu
1               5                   10                  15

Pro Asn His Ser Leu Trp Leu Ala Ser Ala Asn Leu Glu Gly Asp Ala
            20                  25                  30

Leu His Thr Leu Arg Val Thr Leu Val Asp Pro Asn Asn Val Leu Gln
        35                  40                  45

Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val Thr
    50                  55                  60

Cys Asn Asn Glu Asn Ser Val Ile Arg Val Asp Leu Gly Asn Ala Glu
65                  70                  75                  80

Leu Ser Gly His Leu Val Pro Glu Leu Gly Val Leu Lys Asn Leu Gln
                85                  90                  95
```

Glu Leu Tyr Ser Asn Asn Ile Thr Gly Pro Ile Pro Ser Asn Leu Gly
100 105 110

Asn Leu Thr Asn Leu Val Ser Leu Asp Leu Tyr Leu Asn Ser Phe Ser
115 120 125

Gly Pro Ile Pro Glu Ser Leu Gly Lys Leu Ser Lys Leu Arg Phe Leu
130 135 140

Arg Leu Asn Asn Asn Ser Leu Thr Gly Ser Ile Pro Met Ser Leu Thr
145 150 155 160

Asn Ile Thr Thr Leu Gln Val Leu Asp Leu Ser Asn Asn Arg Leu Ser
165 170 175

Gly Ser Val Pro Asp Asn Gly Ser Phe Ser Leu Phe Thr Pro Ile Ser
180 185 190

Phe Ala Asn Asn Leu Asp Leu Cys Gly Pro Val Thr Ser His Pro Cys
195 200 205

Pro Gly Ser Pro Pro Phe Ser Pro Pro Pro Pro Phe Ile Gln Pro Pro
210 215 220

Pro Val Ser Thr Pro Ser Gly Tyr Gly Ile Thr Gly Ala Ile Ala Gly
225 230 235 240

Gly Val Ala Ala Gly Ala Ala Leu Leu Phe Ala Ala Pro Ala Ile Ala
245 250 255

Phe Ala Trp Trp Arg Arg Arg Lys Pro Leu Asp Ile Phe Phe Asp Val
260 265 270

Pro Ala Glu Glu Asp Pro Glu Val His Leu Gly Gln Leu Lys Arg Phe
275 280 285

Ser Leu Arg Glu Leu Gln Val Ala Ser Asp Gly Phe Ser Asn Lys Asn

290 295 300

Ile Leu Gly Arg Gly Gly Phe Gly Lys Val Tyr Lys Gly Arg Leu Ala
305 310 315 320

Asp Gly Thr Leu Val Ala Val Lys Arg Leu Lys Glu Glu Arg Thr Pro
325 330 335

Gly Gly Glu Leu Gln Phe Gln Thr Glu Val Glu Met Ile Ser Met Ala
340 345 350

Val His Arg Asn Leu Leu Arg Leu Arg Gly Phe Cys Met Thr Pro Thr
355 360 365

Glu Arg Leu Leu Val Tyr Pro Tyr Met Ala Asn Gly Ser Val Ala Ser
370 375 380

Cys Leu Arg Glu Arg Pro Pro Ser Gln Pro Pro Leu Asp Trp Pro Thr
385 390 395 400

Arg Lys Arg Ile Ala Leu Gly Ser Ala Arg Gly Leu Ser Tyr Leu His
405 410 415

Asp His Cys Asp Pro Lys Ile Ile His Arg Asp Val Lys Ala Ala Asn
420 425 430

Ile Leu Leu Asp Glu Glu Phe Glu Ala Val Val Gly Asp Phe Gly Leu
435 440 445

Ala Lys Leu Met Asp Tyr Lys Asp Thr His Val Thr Thr Ala Val Arg
450 455 460

Gly Thr Ile Gly His Ile Ala Pro Glu Tyr Leu Ser Thr Gly Lys Ser
465 470 475 480

Ser Glu Lys Thr Asp Val Phe Gly Tyr Gly Ile Met Leu Leu
485 490

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1106 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 142..795

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

```
TCGACCCACG CGTCCGTCCA ACTTCAATAA AGGGGAAACC AACGTAACCC TAATTTTGCT      60

TTCTCCTCTT TGTTCAGAAA ATTTTTCCCTT TACTCTCAAA TTCCTTTTCG ATTTCCCTCT    120

CTTAAACCTC CGAAAGCTCA C ATG GCG TCT CGA AAC TAT CGG TGG GAG CTC      171
                        Met Ala Ser Arg Asn Tyr Arg Trp Glu Leu
                          1               5                  10

TTC GCA GCT TCG TTA ACC CTA ACC TTA GCT TTG ATT CAC CTG GTC GAA      219
Phe Ala Ala Ser Leu Thr Leu Thr Leu Ala Leu Ile His Leu Val Glu
                15                  20                  25

GCA AAC TCC GAA GGA GAT GCT CTC TAC GCT CTT CGC CGG AGT TTG ACA      267
Ala Asn Ser Glu Gly Asp Ala Leu Tyr Ala Leu Arg Arg Ser Leu Thr
            30                  35                  40
```

GAT CCA GAC CAT GTC CTC CAG AGC TGG GAT CCA ACT CTT GTT AAT CCT 315
Asp Pro Asp His Val Leu Gln Ser Trp Asp Pro Thr Leu Val Asn Pro
45 50 55

TGT ACC TGG TTC CAT GTC ACC TGT AAC CAA GAC AAC CGC GTC ACT CGT 363
Cys Thr Trp Phe His Val Thr Cys Asn Gln Asp Asn Arg Val Thr Arg
60 65 70

GTG GAT TTG GGA AAT TCA AAC CTC TCT GGA CAT CTT GCG CCT GAG CTT 411
Val Asp Leu Gly Asn Ser Asn Leu Ser Gly His Leu Ala Pro Glu Leu
75 80 85 90

GGG AAG CTT GAA CAT TTA CAG TAT CTA GAG CTC TAC AAA AAC AAC ATC 459
Gly Lys Leu Glu His Leu Gln Tyr Leu Glu Leu Tyr Lys Asn Asn Ile
95 100 105

CAA GGA ACT ATA CCT TCC GAA CTT GGA AAT CTG AAG AAT CTC ATC AGC 507
Gln Gly Thr Ile Pro Ser Glu Leu Gly Asn Leu Lys Asn Leu Ile Ser
110 115 120

TTG GAT CTG TAC AAC AAC AAT CTT ACA GGG ATA GTT CCC ACT TTC TTG 555
Leu Asp Leu Tyr Asn Asn Asn Leu Thr Gly Ile Val Pro Thr Phe Leu
125 130 135

GGA AAA TTG AAG TCT CTG GTC TTT TTA CGG CTT AAT GAC AAC CGA TTG 603
Gly Lys Leu Lys Ser Leu Val Phe Leu Arg Leu Asn Asp Asn Arg Leu
140 145 150

ACC GGT CCA ATC CTA GAG CAC TCA CGG CAA TCC CAA GCC TTT AAA GTT 651
Thr Gly Pro Ile Leu Glu His Ser Arg Gln Ser Gln Ala Phe Lys Val
155 160 165 170

GTT GAC GTC TCA AGC AAT GAT TTG TGT GGG ACA ATC CCA ACA AAC GGA 699
Val Asp Val Ser Ser Asn Asp Leu Cys Gly Thr Ile Pro Thr Asn Gly
175 180 185

CCC TTT GCT CAC ATT CCT TTA CAG AAC TTT GAG AAC AAC CCG AGA TTG 747

Pro Phe Ala His Ile Pro Leu Gln Asn Phe Glu Asn Asn Pro Arg Leu
190 195 200

GAG GGA CCG GAA TTA CTC GGT CTT GCA AGC TAC GAC ACT AAC TGC ACC 795
Glu Gly Pro Glu Leu Leu Gly Leu Ala Ser Tyr Asp Thr Asn Cys Thr
205 210 215

TGAAACAACT GGCAAAACCT GAAAATGAAG AATTGGGGGG TGACCTTGTA AGAACACTTC 855

ACCACTTTAT CAAATATCAC ATCTATTATG TAATAAGTAT ATATATGTAG TAAAAACAAA 915

AAAAATGAAG AATCGAATCG GTAATATCAT CTGGTCTCAA TTGAGAACTT CGAGGTCTGT 975

ATGTAAAATT TCTAAATGCG ATTTTCGCTT ACTGTAATGT TCGGTTGTGG GATTCTGAGA 1035

AGTAACATTT GTATTGGTAT GGTATCAAGT TGTTCTGCCT TGTCTGCAAA AAAAAAAAAA 1095

AAAAAAAAAA A 1106

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 218 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

Met Ala Ser Arg Asn Tyr Arg Trp Glu Leu Phe Ala Ala Ser Leu Thr
1 5 10 15

Leu Thr Leu Ala Leu Ile His Leu Val Glu Ala Asn Ser Glu Gly Asp
20 25 30

Ala Leu Tyr Ala Leu Arg Arg Ser Leu Thr Asp Pro Asp His Val Leu
35 40 45

Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val
50 55 60

Thr Cys Asn Gln Asp Asn Arg Val Thr Arg Val Asp Leu Gly Asn Ser
65 70 75 80

Asn Leu Ser Gly His Leu Ala Pro Glu Leu Gly Lys Leu Glu His Leu
85 90 95

Gln Tyr Leu Glu Leu Tyr Lys Asn Asn Ile Gln Gly Thr Ile Pro Ser
100 105 110

Glu Leu Gly Asn Leu Lys Asn Leu Ile Ser Leu Asp Leu Tyr Asn Asn
115 120 125

Asn Leu Thr Gly Ile Val Pro Thr Phe Leu Gly Lys Leu Lys Ser Leu
130 135 140

Val Phe Leu Arg Leu Asn Asp Asn Arg Leu Thr Gly Pro Ile Leu Glu
145 150 155 160

His Ser Arg Gln Ser Gln Ala Phe Lys Val Val Asp Val Ser Ser Asn
165 170 175

Asp Leu Cys Gly Thr Ile Pro Thr Asn Gly Pro Phe Ala His Ile Pro
180 185 190

Leu Gln Asn Phe Glu Asn Asn Pro Arg Leu Glu Gly Pro Glu Leu Leu
195 200 205

Gly Leu Ala Ser Tyr Asp Thr Asn Cys Thr
210 215

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 981 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 104..757

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

```
AGTGTGAGTA ATTTAGTTTG CTTTCTCCTC TTTGTTCAGA AAATTTTCCC TTTACTCTCA        60

AATTCCTTTT CGATTTCCCT CTCTTAAACC TCCGAAAGCT CAC ATG GCG TCT CGA       115
                                                Met Ala Ser Arg
                                                  1

AAC TAT CGG TGG GAG CTC TTC GCA GCT TCG TTA ACC CTA ACC TTA GCT       163
Asn Tyr Arg Trp Glu Leu Phe Ala Ala Ser Leu Thr Leu Thr Leu Ala
  5                  10                  15                  20

TTG ATT CAC CTG GTC GAA GCA AAC TCC GAA GGA GAT GCT CTC TAC GCT       211
Leu Ile His Leu Val Glu Ala Asn Ser Glu Gly Asp Ala Leu Tyr Ala
                 25                  30                  35

CTT CGC CGG AGT TTG ACA GAT CCA GAC CAT GTC CTC CAG AGC TGG GAT       259
Leu Arg Arg Ser Leu Thr Asp Pro Asp His Val Leu Gln Ser Trp Asp
             40                  45                  50
```

CCA ACT CTT GTT AAT CCT TGT ACC TGG TTC CAT GTC ACC TGT AAC CAA 307

Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val Thr Cys Asn Gln

55 60 65

GAC AAC CGC GTC ACT CGT GTG GAT TTG GGA AAT TCA AAC CTC TCT GGA 355

Asp Asn Arg Val Thr Arg Val Asp Leu Gly Asn Ser Asn Leu Ser Gly

70 75 80

CAT CTT GCG CCT GAG CTT GGG AAG CTT GAA CAT TTA CAG TAT CTA GAG 403

His Leu Ala Pro Glu Leu Gly Lys Leu Glu His Leu Gln Tyr Leu Glu

85 90 95 100

CTC TAC AAA AAC AAC ATC CAA GGA ACT ATA CCT TCC GAA CTT GGA AAT 451

Leu Tyr Lys Asn Asn Ile Gln Gly Thr Ile Pro Ser Glu Leu Gly Asn

105 110 115

CTG AAG AAT CTC ATC AGC TTG GAT CTG TAC AAC AAC AAT CTT ACA GGG 499

Leu Lys Asn Leu Ile Ser Leu Asp Leu Tyr Asn Asn Asn Leu Thr Gly

120 125 130

ATA GTT CCC ACT TCT TTG GGA AAA TTG AAG TCT CTG GTC TTT TTA CGG 547

Ile Val Pro Thr Ser Leu Gly Lys Leu Lys Ser Leu Val Phe Leu Arg

135 140 145

CTT AAT GAC AAC CGA TTG ACC GGT CCA ATC CCT AGA GCA CTC ACG GCA 595

Leu Asn Asp Asn Arg Leu Thr Gly Pro Ile Pro Arg Ala Leu Thr Ala

150 155 160

ATC CCA AGC CTT AAA GTT GTT GAC GTC TCA AGC AAT GAT TTG TGT GGA 643

Ile Pro Ser Leu Lys Val Val Asp Val Ser Ser Asn Asp Leu Cys Gly

165 170 175 180

ACA ATC CCA ACA AAC GGA CCC TTT GCT CAC ATT CCT TTA CAG AAC TTT 691

Thr Ile Pro Thr Asn Gly Pro Phe Ala His Ile Pro Leu Gln Asn Phe

185 190 195

GAG AAC AAC CCG AGA TTG GAG GGA CCG GAA TTA CTC GGT CTT GCA AGC 739

Glu Asn Asn Pro Arg Leu Glu Gly Pro Glu Leu Leu Gly Leu Ala Ser
200 205 210

TAC GAC ACT AAC TGC ACC TGAAACAACT GGCAAAACCT GAAAATGAAG 787
Tyr Asp Thr Asn Cys Thr
215

AATTGGGGGG TGACCTTGTA AGAACACTTC ACCACTTTAT CAAATATCAC ATCTATTATG 847

TAATAAGTAT ATATATGTAG TAAAAACAAA AAAAATGAAG AATCGAATCG GTAATATCAT 907

CTGGTCTCAA TTGAGAACTT CGAGGTCTGT ATGTAAAATT TCTAAATGCG ATTTTTCGCCT 967

AAATTACTCA CACT 981

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 218 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Met Ala Ser Arg Asn Tyr Arg Trp Glu Leu Phe Ala Ala Ser Leu Thr
1 5 10 15

Leu Thr Leu Ala Leu Ile His Leu Val Glu Ala Asn Ser Glu Gly Asp
20 25 30

Ala Leu Tyr Ala Leu Arg Arg Ser Leu Thr Asp Pro Asp His Val Leu
35 40 45

Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val

50 55 60

Thr Cys Asn Gln Asp Asn Arg Val Thr Arg Val Asp Leu Gly Asn Ser
65 70 75 80

Asn Leu Ser Gly His Leu Ala Pro Glu Leu Gly Lys Leu Glu His Leu
85 90 95

Gln Tyr Leu Glu Leu Tyr Lys Asn Asn Ile Gln Gly Thr Ile Pro Ser
100 105 110

Glu Leu Gly Asn Leu Lys Asn Leu Ile Ser Leu Asp Leu Tyr Asn Asn
115 120 125

Asn Leu Thr Gly Ile Val Pro Thr Ser Leu Gly Lys Leu Lys Ser Leu
130 135 140

Val Phe Leu Arg Leu Asn Asp Asn Arg Leu Thr Gly Pro Ile Pro Arg
145 150 155 160

Ala Leu Thr Ala Ile Pro Ser Leu Lys Val Val Asp Val Ser Ser Asn
165 170 175

Asp Leu Cys Gly Thr Ile Pro Thr Asn Gly Pro Phe Ala His Ile Pro
180 185 190

Leu Gln Asn Phe Glu Asn Asn Pro Arg Leu Glu Gly Pro Glu Leu Leu
195 200 205

Gly Leu Ala Ser Tyr Asp Thr Asn Cys Thr
210 215

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 789 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..661

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

```
T CGA CCC ACG CGT CCG CGA AAC TAT CGG TGG GAG CTC TTC GCA GCT        46
  Arg Pro Thr Arg Pro Arg Asn Tyr Arg Trp Glu Leu Phe Ala Ala
    1               5                  10                  15

TCG TTA ATC CTA ACC TTA GCT TTG ATT CAC CTG GTC GAA GCA AAC TCC      94
Ser Leu Ile Leu Thr Leu Ala Leu Ile His Leu Val Glu Ala Asn Ser
                 20                  25                  30

GAA GGA GAT GCT CTT TAC GCT CTT CGC CGG AGT TTA ACA GAT CCG GAC     142
Glu Gly Asp Ala Leu Tyr Ala Leu Arg Arg Ser Leu Thr Asp Pro Asp
             35                  40                  45

CAT GTT CTC CAG AGC TGG GAT CCA ACT CTT GTT AAT CCT TGT ACC TGG     190
His Val Leu Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp
         50                  55                  60

TTC CAT GTC ACC TGT AAC CAA GAC AAC CGC GTC ACT CGT GTG GAT TTG     238
Phe His Val Thr Cys Asn Gln Asp Asn Arg Val Thr Arg Val Asp Leu
     65                  70                  75

GGG AAT TCA AAC CTC TCT GGA CAT CTT GCG CCT GAG CTT GGG AAG CTT     286
Gly Asn Ser Asn Leu Ser Gly His Leu Ala Pro Glu Leu Gly Lys Leu
```

80 85 90 95

GAA CAT TTA CAG TAT CTA GAG CTC TAC AAA AAC AAC ATC CAA GGA ACT 334
Glu His Leu Gln Tyr Leu Glu Leu Tyr Lys Asn Asn Ile Gln Gly Thr
100 105 110

ATA CCT TCC GAA CTT GGA AAT CTG AAG AAT CTC ATC AGC TTG GAT CTG 382
Ile Pro Ser Glu Leu Gly Asn Leu Lys Asn Leu Ile Ser Leu Asp Leu
115 120 125

TAC AAC AAC AAT CTT ACA GGG ATA GTT CCC ACT TCT TTG GGA AAA TTG 430
Tyr Asn Asn Asn Leu Thr Gly Ile Val Pro Thr Ser Leu Gly Lys Leu
130 135 140

AAG TCT CTG GTC TTT TTA CGG CTT AAT GAC AAC CGA TTG ACG GGG CCA 478
Lys Ser Leu Val Phe Leu Arg Leu Asn Asp Asn Arg Leu Thr Gly Pro
145 150 155

ATC CCT AGA GCA CTC ACT GCA ATC CCA AGC CTT AAA GTT GTT GAT GTC 526
Ile Pro Arg Ala Leu Thr Ala Ile Pro Ser Leu Lys Val Val Asp Val
160 165 170 175

TCA AGC AAT GAT TTG TGT GGA ACA ATC CCA ACA AAC GGA CCT TTT GCT 574
Ser Ser Asn Asp Leu Cys Gly Thr Ile Pro Thr Asn Gly Pro Phe Ala
180 185 190

CAC ATT CCT TTA CAG AAC TTT GAG AAC AAC CCG AGG TTG GAG GGA CCG 622
His Ile Pro Leu Gln Asn Phe Glu Asn Asn Pro Arg Leu Glu Gly Pro
195 200 205

GAA TTA CTC GGT CTT GCA AGC TAC GAC ACT AAC TGC ACC TGAAAAAATT 671
Glu Leu Leu Gly Leu Ala Ser Tyr Asp Thr Asn Cys Thr
210 215 220

GGCAAAACCT GAAAATGAAG AATTGGGGGG TGACCTTGTA AGAACACTTC ACCACTTTAT 731

CAAATATCAC ATCTACTATG TAATAAGTAT ATATATGTAG TCCAAAAAAA AAAAAAAA 789

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 220 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Arg Pro Thr Arg Pro Arg Asn Tyr Arg Trp Glu Leu Phe Ala Ala Ser
1 5 10 15

Leu Ile Leu Thr Leu Ala Leu Ile His Leu Val Glu Ala Asn Ser Glu
20 25 30

Gly Asp Ala Leu Tyr Ala Leu Arg Arg Ser Leu Thr Asp Pro Asp His
35 40 45

Val Leu Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe
50 55 60

His Val Thr Cys Asn Gln Asp Asn Arg Val Thr Arg Val Asp Leu Gly
65 70 75 80

Asn Ser Asn Leu Ser Gly His Leu Ala Pro Glu Leu Gly Lys Leu Glu
85 90 95

His Leu Gln Tyr Leu Glu Leu Tyr Lys Asn Asn Ile Gln Gly Thr Ile
100 105 110

Pro Ser Glu Leu Gly Asn Leu Lys Asn Leu Ile Ser Leu Asp Leu Tyr
115 120 125

Asn Asn Asn Leu Thr Gly Ile Val Pro Thr Ser Leu Gly Lys Leu Lys
130 135 140

Ser Leu Val Phe Leu Arg Leu Asn Asp Asn Arg Leu Thr Gly Pro Ile
145 150 155 160

Pro Arg Ala Leu Thr Ala Ile Pro Ser Leu Lys Val Val Asp Val Ser
165 170 175

Ser Asn Asp Leu Cys Gly Thr Ile Pro Thr Asn Gly Pro Phe Ala His
180 185 190

Ile Pro Leu Gln Asn Phe Glu Asn Asn Pro Arg Leu Glu Gly Pro Glu
195 200 205

Leu Leu Gly Leu Ala Ser Tyr Asp Thr Asn Cys Thr
210 215 220

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 894 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..675

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GGA CCG ATT CAA GCC TCC GAA GGG GAC GCT CTT CAC GCG CTT CGC CGG 48
Gly Pro Ile Gln Ala Ser Glu Gly Asp Ala Leu His Ala Leu Arg Arg
1 5 10 15

AGC TTA TCA GAT CCA GAC AAT GTT GTT CAG AGT TGG GAT CCA ACT CTT 96
Ser Leu Ser Asp Pro Asp Asn Val Val Gln Ser Trp Asp Pro Thr Leu
20 25 30

GTT AAT CCT TGT ACT TGG TTT CAT GTC ACT TGT AAT CAA CAC CAT CAA 144
Val Asn Pro Cys Thr Trp Phe His Val Thr Cys Asn Gln His His Gln
35 40 45

GTC ACT CGT CTG GAT TTG GGG AAT TCA AAC TTA TCT GGA CAT CTA GTA 192
Val Thr Arg Leu Asp Leu Gly Asn Ser Asn Leu Ser Gly His Leu Val
50 55 60

CCT GAA CTT GGG AAG CTT GAA CAT TTA CAA TAT CTG TAT GGA ATC ATC 240
Pro Glu Leu Gly Lys Leu Glu His Leu Gln Tyr Leu Tyr Gly Ile Ile
65 70 75 80

ACT CTT TTG CCT TTT GAT TAT CTG AAA ACA TTT ACA TTA TCA GTC ACA 288
Thr Leu Leu Pro Phe Asp Tyr Leu Lys Thr Phe Thr Leu Ser Val Thr
85 90 95

CAT ATA ACA TTT TGC TTT GAG TCA TAT AGT GAA CTC TAC AAA AAC GAG 336
His Ile Thr Phe Cys Phe Glu Ser Tyr Ser Glu Leu Tyr Lys Asn Glu
100 105 110

ATT CAA GGA ACT ATA CCT TCT GAG CTT GGA AAT CTG AAG AGT CTA ATC 384
Ile Gln Gly Thr Ile Pro Ser Glu Leu Gly Asn Leu Lys Ser Leu Ile
115 120 125

AGT TTG GAT CTG TAC AAC AAC AAT CTC ACC GGG AAA ATC CCA TCT TCT 432
Ser Leu Asp Leu Tyr Asn Asn Asn Leu Thr Gly Lys Ile Pro Ser Ser
130 135 140

```
TTG GGA AAA TTG AAG TCA CTT GTT TTT TTG CGG CTT AAC GAA AAC CGA      480
Leu Gly Lys Leu Lys Ser Leu Val Phe Leu Arg Leu Asn Glu Asn Arg
145                 150                 155                 160

TTG ACC GGT CCT ATT CCT AGA GAA CTC ACA GTT ATT TCA AGC CTT AAA      528
Leu Thr Gly Pro Ile Pro Arg Glu Leu Thr Val Ile Ser Ser Leu Lys
                165                 170                 175

GTT GTT GAT GTC TCA GGG AAT GAT TTG TGT GGA ACA ATT CCA GTA GAA      576
Val Val Asp Val Ser Gly Asn Asp Leu Cys Gly Thr Ile Pro Val Glu
            180                 185                 190

GGA CCT TTT GAA CAC ATT CCT ATG CAA AAC TTT GAG AAC AAC CTG AGA      624
Gly Pro Phe Glu His Ile Pro Met Gln Asn Phe Glu Asn Asn Leu Arg
        195                 200                 205

TTG GAG GGA CCA GAA CTA CTA GGT CTT GCG AGC TAT GAC ACC AAT TGC      672
Leu Glu Gly Pro Glu Leu Leu Gly Leu Ala Ser Tyr Asp Thr Asn Cys
    210                 215                 220

ACT TAAAAAGAAG TTGAAGAACC TATAAAGAAG AATGTTAGGT GACCTTGTAA           725
Thr
225

GAACTCTGTA CCAAGTGTTT GTAAATCTAT ATAGAGCCTT GTTTCATGTT ATATATGAAA    785

GCTTTGAGAG ACAGTAACTT GCAATGTATT GGTATTGGTA GAAAAAGTTG AAATGAGAAT    845

TGCTTTGTAA TTGGATTTGT GTTTCTTATG TAACTTGAAT TTCTTATTA                894
```

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 225 amino acids
  (B) TYPE: amino acid
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Gly Pro Ile Gln Ala Ser Glu Gly Asp Ala Leu His Ala Leu Arg Arg
1 5 10 15

Ser Leu Ser Asp Pro Asp Asn Val Val Gln Ser Trp Asp Pro Thr Leu
20 25 30

Val Asn Pro Cys Thr Trp Phe His Val Thr Cys Asn Gln His His Gln
35 40 45

Val Thr Arg Leu Asp Leu Gly Asn Ser Asn Leu Ser Gly His Leu Val
50 55 60

Pro Glu Leu Gly Lys Leu Glu His Leu Gln Tyr Leu Tyr Gly Ile Ile
65 70 75 80

Thr Leu Leu Pro Phe Asp Tyr Leu Lys Thr Phe Thr Leu Ser Val Thr
85 90 95

His Ile Thr Phe Cys Phe Glu Ser Tyr Ser Glu Leu Tyr Lys Asn Glu
100 105 110

Ile Gln Gly Thr Ile Pro Ser Glu Leu Gly Asn Leu Lys Ser Leu Ile
115 120 125

Ser Leu Asp Leu Tyr Asn Asn Asn Leu Thr Gly Lys Ile Pro Ser Ser
130 135 140

Leu Gly Lys Leu Lys Ser Leu Val Phe Leu Arg Leu Asn Glu Asn Arg
145 150 155 160

Leu Thr Gly Pro Ile Pro Arg Glu Leu Thr Val Ile Ser Ser Leu Lys
165 170 175

Val Val Asp Val Ser Gly Asn Asp Leu Cys Gly Thr Ile Pro Val Glu
180 185 190

Gly Pro Phe Glu His Ile Pro Met Gln Asn Phe Glu Asn Asn Leu Arg
195 200 205

Leu Glu Gly Pro Glu Leu Leu Gly Leu Ala Ser Tyr Asp Thr Asn Cys
210 215 220

Thr
225

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1063 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 106..759

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

TCGACCCACG CGTCCGACGA AACCCTAATT TTGCTTCCTC ATCTTGTTCA GAAAATTACT 60

CAAATTCCTA TTAGATTACT CTCTCTTCGA CCTCCGATAG CTCAC ATG GCG TCT 114
Met Ala Ser

1

CGA AAC TAT CGG TGG GAG CTC TTC GCA GCT TCG TTA ATC CTA ACC TTA 162

Arg Asn Tyr Arg Trp Glu Leu Phe Ala Ala Ser Leu Ile Leu Thr Leu

5 10 15

GCT TTG ATT CAC CTG GTC GAA GCA AAC TCC GAA GGA GAT GCT CTT TAC 210

Ala Leu Ile His Leu Val Glu Ala Asn Ser Glu Gly Asp Ala Leu Tyr

20 25 30 35

GCT CTT CGC CGG AGT TTA ACA GAT CCG GAC CAT GTT CTC CAG AGC TGG 258

Ala Leu Arg Arg Ser Leu Thr Asp Pro Asp His Val Leu Gln Ser Trp

40 45 50

GAT CCA ACT CTT GTT AAT CCT TGT ACC TGG TTC CAT GTC ACC TGT AAC 306

Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val Thr Cys Asn

55 60 65

CAA GAC AAC CGC GTC ACT CGT GTG GAT TTG GGG AAT TCA AAC CTC TCT 354

Gln Asp Asn Arg Val Thr Arg Val Asp Leu Gly Asn Ser Asn Leu Ser

70 75 80

GGA CAT CTT GCG CCT GAG CTT GGG AAG CTT GAA CAT TTA CAG TAT CTA 402

Gly His Leu Ala Pro Glu Leu Gly Lys Leu Glu His Leu Gln Tyr Leu

85 90 95

GAG CTC TAC AAA AAC AAC ATC CAA GGA ACT ATA CCT TCC GAA CTT GGA 450

Glu Leu Tyr Lys Asn Asn Ile Gln Gly Thr Ile Pro Ser Glu Leu Gly

100 105 110 115

AAT CTG AAG AAT CTC ATC AGC TTG GAT CTG TAC AAC AAC AAT CTT ACA 498

Asn Leu Lys Asn Leu Ile Ser Leu Asp Leu Tyr Asn Asn Asn Leu Thr

120 125 130

GGG ATA GTT CCC ACT TCT TTG GGA AAA TTG AAG TCT CTG GTC TTT TTA 546

Gly Ile Val Pro Thr Ser Leu Gly Lys Leu Lys Ser Leu Val Phe Leu

135 140 145

CGG CTT AAT GAC AAC CGA TTG ACG GGG CCA ATC CCT AGA GCA CTC ACT 594
Arg Leu Asn Asp Asn Arg Leu Thr Gly Pro Ile Pro Arg Ala Leu Thr
150 155 160

GCA ATC CCA AGC CTT AAA GTT GTT GAT GTC TCA AGC AAT GAT TTG TGT 642
Ala Ile Pro Ser Leu Lys Val Val Asp Val Ser Ser Asn Asp Leu Cys
165 170 175

GGA ACA ATC CCA ACA AAC GGA CCT TTT GCT CAC ATT CCT TTA CAG AAC 690
Gly Thr Ile Pro Thr Asn Gly Pro Phe Ala His Ile Pro Leu Gln Asn
180 185 190 195

TTT GAG AAC AAC CCG AGG TTG GAG GGA CCG GAA TTA CTC GGT CTT GCA 738
Phe Glu Asn Asn Pro Arg Leu Glu Gly Pro Glu Leu Leu Gly Leu Ala
200 205 210

AGC TAC GAC ACT AAC TGC ACC TGAAAAAATT GGCAAAACCT GAAAATGAAG 789
Ser Tyr Asp Thr Asn Cys Thr
215

AATTGGGGGG TGACCTTGTA AGAACACTTC ACCACTTTAT CAAATATCAC ATCTACTATG 849

TAATAAGTAT ATATATGTAG TCCAAAAAAA AAATGAAGAA TCGAATCAGT AATATCATCT 909

GGTCTCAATT GAGAACTTTG AGGTCTGTGT ATGTAAAATT TCTAAATGCG ACTTTCGCGT 969

ACTGTAATGT TCGGTTGTGG GATTCTGAGA AGTAACATTT GTATTGGTAT GGTATCAAGT 1029

TGTTCTGCCT TGTCTGCAAA AAAAAAAAAA AAAA 1063

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 218 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

Met Ala Ser Arg Asn Tyr Arg Trp Glu Leu Phe Ala Ala Ser Leu Ile
1 5 10 15

Leu Thr Leu Ala Leu Ile His Leu Val Glu Ala Asn Ser Glu Gly Asp
20 25 30

Ala Leu Tyr Ala Leu Arg Arg Ser Leu Thr Asp Pro Asp His Val Leu
35 40 45

Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val
50 55 60

Thr Cys Asn Gln Asp Asn Arg Val Thr Arg Val Asp Leu Gly Asn Ser
65 70 75 80

Asn Leu Ser Gly His Leu Ala Pro Glu Leu Gly Lys Leu Glu His Leu
85 90 95

Gln Tyr Leu Glu Leu Tyr Lys Asn Asn Ile Gln Gly Thr Ile Pro Ser
100 105 110

Glu Leu Gly Asn Leu Lys Asn Leu Ile Ser Leu Asp Leu Tyr Asn Asn
115 120 125

Asn Leu Thr Gly Ile Val Pro Thr Ser Leu Gly Lys Leu Lys Ser Leu
130 135 140

Val Phe Leu Arg Leu Asn Asp Asn Arg Leu Thr Gly Pro Ile Pro Arg
145 150 155 160

Ala Leu Thr Ala Ile Pro Ser Leu Lys Val Val Asp Val Ser Ser Asn

165 170 175

Asp Leu Cys Gly Thr Ile Pro Thr Asn Gly Pro Phe Ala His Ile Pro
180 185 190

Leu Gln Asn Phe Glu Asn Asn Pro Arg Leu Glu Gly Pro Glu Leu Leu
195 200 205

Gly Leu Ala Ser Tyr Asp Thr Asn Cys Thr
210 215

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2089 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(vii) IMMEDIATE SOURCE:

(B) CLONE: SERK gene cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 195..2069

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

```
GGATTTTTAT TTTATTTTTT ACTCTTTGTT TGTTTTAATG CTAATGGGTT TTTAAAAGGG      60

TTATCGAAAA AATGAGTGAG TTTGTGTTGA GGTTGTCTCT GTAAAGTGTT AATGGTGGTG     120

ATTTTCGGAA GTTAGGGTTT TCTCGGATCT GAAGAGATCA AATCAAGATT CGAAATTTAC     180

CATTGTTGTT TGAA ATG GAG TCG AGT TAT GTG GTG TTT ATC TTA CTT TCA       230
                Met Glu Ser Ser Tyr Val Val Phe Ile Leu Leu Ser
                  1               5                  10

CTG ATC TTA CTT CCG AAT CAT TCA CTG TGG CTT GCT TCT GCT AAT TTG       278
Leu Ile Leu Leu Pro Asn His Ser Leu Trp Leu Ala Ser Ala Asn Leu
         15                  20                  25

GAA GGT GAT GCT TTG CAT ACT TTG AGG GTT ACT CTA GTT GAT CCA AAC       326
Glu Gly Asp Ala Leu His Thr Leu Arg Val Thr Leu Val Asp Pro Asn
     30                  35                  40

AAT GTC TTG CAG AGC TGG GAT CCT ACG CTA GTG AAT CCT TGC ACA TGG       374
Asn Val Leu Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp
 45                  50                  55                  60

TTC CAT GTC ACT TGC AAC AAC GAG AAC AGT GTC ATA AGA GTT GAT TTG       422
Phe His Val Thr Cys Asn Asn Glu Asn Ser Val Ile Arg Val Asp Leu
                 65                  70                  75

GGG AAT GCA GAG TTA TCT GGC CAT TTA GTT CCA GAG CTT GGT GTG CTC       470
Gly Asn Ala Glu Leu Ser Gly His Leu Val Pro Glu Leu Gly Val Leu
             80                  85                  90

AAG AAT TTG CAG TAT TTG GAG CTT TAC AGT AAC AAC ATA ACT GGC CCG       518
Lys Asn Leu Gln Tyr Leu Glu Leu Tyr Ser Asn Asn Ile Thr Gly Pro
         95                 100                 105

ATT CCT AGT AAT CTT GGA AAT CTG ACA AAC TTA GTG AGT TTG GAT CTT       566
```

Ile Pro Ser Asn Leu Gly Asn Leu Thr Asn Leu Val Ser Leu Asp Leu
110 115 120

TAC TTA AAC AGC TTC TCC GGT CCT ATT CCG GAA TCA TTG GGA AAG CTT 614
Tyr Leu Asn Ser Phe Ser Gly Pro Ile Pro Glu Ser Leu Gly Lys Leu
125 130 135 140

TCA AAG CTG AGA TTT CTC CGG CTT AAC AAC AAC AGT CTC ACT GGG TCA 662
Ser Lys Leu Arg Phe Leu Arg Leu Asn Asn Asn Ser Leu Thr Gly Ser
145 150 155

ATT CCT ATG TCA CTG ACC AAT ATT ACT ACC CTT CAA GTG TTA GAT CTA 710
Ile Pro Met Ser Leu Thr Asn Ile Thr Thr Leu Gln Val Leu Asp Leu
160 165 170

TCA AAT AAC AGA CTC TCT GGT TCA GTT CCT GAC AAT GGC TCC TTC TCA 758
Ser Asn Asn Arg Leu Ser Gly Ser Val Pro Asp Asn Gly Ser Phe Ser
175 180 185

CTC TTC ACA CCC ATC AGT TTT GCT AAT AAC TTA GAC CTA TGT GGA CCT 806
Leu Phe Thr Pro Ile Ser Phe Ala Asn Asn Leu Asp Leu Cys Gly Pro
190 195 200

GTT ACA AGT CAC CCA TGT CCT GGA TCT CCC CCG TTT TCT CCT CCA CCA 854
Val Thr Ser His Pro Cys Pro Gly Ser Pro Pro Phe Ser Pro Pro Pro
205 210 215 220

CCT TTT ATT CAA CCT CCC CCA GTT TCC ACC CCG AGT GGG TAT GGT ATA 902
Pro Phe Ile Gln Pro Pro Pro Val Ser Thr Pro Ser Gly Tyr Gly Ile
225 230 235

ACT GGA GCA ATA GCT GGT GGA GTT GCT GCA GGT GCT GCT TTG CCC TTT 950
Thr Gly Ala Ile Ala Gly Gly Val Ala Ala Gly Ala Ala Leu Pro Phe
240 245 250

GCT GCT CCT GCA ATA GCC TTT GCT TGG TGG CGA CGA AGA AGC CCA CTA 998
Ala Ala Pro Ala Ile Ala Phe Ala Trp Trp Arg Arg Arg Ser Pro Leu

255 260 265

GAT ATT TTC TTC GAT GTC CCT GCC GAA GAA GAT CCA GAA GTT CAT CTG 1046
Asp Ile Phe Phe Asp Val Pro Ala Glu Glu Asp Pro Glu Val His Leu
270 275 280

GGA CAG CTC AAG AGG TTT TCT TTG CGG GAG CTA CAA GTG GCG AGT GAT 1094
Gly Gln Leu Lys Arg Phe Ser Leu Arg Glu Leu Gln Val Ala Ser Asp
285 290 295 300

GGG TTT AGT AAC AAG AAC ATT TTG GGC AGA GGT GGG TTT GGG AAA GTC 1142
Gly Phe Ser Asn Lys Asn Ile Leu Gly Arg Gly Gly Phe Gly Lys Val
305 310 315

TAC AAG GGA CGC TTG GCA GAC GGA ACT CTT GTT GCT GTC AAG AGA CTG 1190
Tyr Lys Gly Arg Leu Ala Asp Gly Thr Leu Val Ala Val Lys Arg Leu
320 325 330

AAG GAA GAG CGA ACT CCA GGT GGA GAG CTC CAG TTT CAA ACA GAA GTA 1238
Lys Glu Glu Arg Thr Pro Gly Gly Glu Leu Gln Phe Gln Thr Glu Val
335 340 345

GAG ATG ATA AGT ATG GCA GTT CAT CGA AAC CTG TTG AGA TTA CGA GGT 1286
Glu Met Ile Ser Met Ala Val His Arg Asn Leu Leu Arg Leu Arg Gly
350 355 360

TTC TGT ATG ACA CCG ACC GAG AGA TTG CTT GTG TAT CCT TAC ATG GCC 1334
Phe Cys Met Thr Pro Thr Glu Arg Leu Leu Val Tyr Pro Tyr Met Ala
365 370 375 380

AAT GGA AGT GTT GCT TCG TGT CTC AGA GAG AGG CCA CCG TCA CAA CCT 1382
Asn Gly Ser Val Ala Ser Cys Leu Arg Glu Arg Pro Pro Ser Gln Pro
385 390 395

CCG CTT GAT TGG CCA ACG CGG AAG AGA ATC GCG CTA GGC TCA GCT CGA 1430
Pro Leu Asp Trp Pro Thr Arg Lys Arg Ile Ala Leu Gly Ser Ala Arg
400 405 410

```
GGT TTG TCT TAC CTA CAT GAT CAC TGC GAT CCG AAG ATC ATT CAC CGT      1478
Gly Leu Ser Tyr Leu His Asp His Cys Asp Pro Lys Ile Ile His Arg
        415                 420                 425

GAC GTA AAA GCA GCA AAC ATC CTC TTA GAC GAA GAA TTC GAA GCG GTT      1526
Asp Val Lys Ala Ala Asn Ile Leu Leu Asp Glu Glu Phe Glu Ala Val
    430                 435                 440

GTT GGA GAT TTC GGG TTG GCA AAG CTT ATG GAC TAT AAA GAC ACT CAC      1574
Val Gly Asp Phe Gly Leu Ala Lys Leu Met Asp Tyr Lys Asp Thr His
445                 450                 455                 460

GTG ACA ACA GCA GTC CGT GGC ACC ATC GGT CAC ATC GCT CCA GAA TAT      1622
Val Thr Thr Ala Val Arg Gly Thr Ile Gly His Ile Ala Pro Glu Tyr
                465                 470                 475

CTC TCA ACC GGA AAA TCT TCA GAG AAA ACC GAC GTT TTC GGA TAC GGA      1670
Leu Ser Thr Gly Lys Ser Ser Glu Lys Thr Asp Val Phe Gly Tyr Gly
            480                 485                 490

ATC ATG CTT CTA GAA CTA ATC ACA GGA CAA AGA GCT TTC GAT CTC GCT      1718
Ile Met Leu Leu Glu Leu Ile Thr Gly Gln Arg Ala Phe Asp Leu Ala
        495                 500                 505

CGG CTA GCT AAC GAC GAC GAC GTC ATG TTA CTT GAC TGG GTG AAA GGA      1766
Arg Leu Ala Asn Asp Asp Asp Val Met Leu Leu Asp Trp Val Lys Gly
    510                 515                 520

TTG TTG AAG GAG AAG AAG CTA GAG ATG TTA GTG GAT CCA GAT CTT CAA      1814
Leu Leu Lys Glu Lys Lys Leu Glu Met Leu Val Asp Pro Asp Leu Gln
525                 530                 535                 540

ACA AAC TAC GAG GAG AGA GAA CTG GAA CAA GTG ATA CAA GTG GCG TTG      1862
Thr Asn Tyr Glu Glu Arg Glu Leu Glu Gln Val Ile Gln Val Ala Leu
                545                 550                 555
```

WEST

CTA TGC ACG CAA GGA TCA CCA ATG GAA AGA CCA AAG ATG TCT GAA GTT 1910
Leu Cys Thr Gln Gly Ser Pro Met Glu Arg Pro Lys Met Ser Glu Val
560 565 570

GTA AGG ATG CTG GAA GGA GAT GGG CTT GCG GAG AAA TGG GAC GAA TGG 1958
Val Arg Met Leu Glu Gly Asp Gly Leu Ala Glu Lys Trp Asp Glu Trp
575 580 585

CAA AAA GTT GAG ATT TTG AGG GAA GAG ATT GAT TTG AGT CCT AAT CCT 2006
Gln Lys Val Glu Ile Leu Arg Glu Glu Ile Asp Leu Ser Pro Asn Pro
590 595 600

AAC TCT GAT TGG ATT CTT GAT TCT ACT TAC AAT TTG CAC GCC GTT GAG 2054
Asn Ser Asp Trp Ile Leu Asp Ser Thr Tyr Asn Leu His Ala Val Glu
605 610 615 620

TTA TCT GGT CCA AGG TAAAAAAAAA AAAAAAAAAA 2089
Leu Ser Gly Pro Arg
625

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 625 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

Met Glu Ser Ser Tyr Val Val Phe Ile Leu Leu Ser Leu Ile Leu Leu
1 5 10 15

Pro Asn His Ser Leu Trp Leu Ala Ser Ala Asn Leu Glu Gly Asp Ala
20 25 30

Leu His Thr Leu Arg Val Thr Leu Val Asp Pro Asn Asn Val Leu Gln
35 40 45

Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val Thr
50 55 60

Cys Asn Asn Glu Asn Ser Val Ile Arg Val Asp Leu Gly Asn Ala Glu
65 70 75 80

Leu Ser Gly His Leu Val Pro Glu Leu Gly Val Leu Lys Asn Leu Gln
85 90 95

Tyr Leu Glu Leu Tyr Ser Asn Asn Ile Thr Gly Pro Ile Pro Ser Asn
100 105 110

Leu Gly Asn Leu Thr Asn Leu Val Ser Leu Asp Leu Tyr Leu Asn Ser
115 120 125

Phe Ser Gly Pro Ile Pro Glu Ser Leu Gly Lys Leu Ser Lys Leu Arg
130 135 140

Phe Leu Arg Leu Asn Asn Asn Ser Leu Thr Gly Ser Ile Pro Met Ser
145 150 155 160

Leu Thr Asn Ile Thr Thr Leu Gln Val Leu Asp Leu Ser Asn Asn Arg
165 170 175

Leu Ser Gly Ser Val Pro Asp Asn Gly Ser Phe Ser Leu Phe Thr Pro
180 185 190

Ile Ser Phe Ala Asn Asn Leu Asp Leu Cys Gly Pro Val Thr Ser His
195 200 205

Pro Cys Pro Gly Ser Pro Pro Phe Ser Pro Pro Pro Pro Phe Ile Gln
210 215 220

Pro Pro Pro Val Ser Thr Pro Ser Gly Tyr Gly Ile Thr Gly Ala Ile
225 230 235 240

Ala Gly Gly Val Ala Ala Gly Ala Ala Leu Pro Phe Ala Ala Pro Ala
245 250 255

Ile Ala Phe Ala Trp Trp Arg Arg Arg Ser Pro Leu Asp Ile Phe Phe
260 265 270

Asp Val Pro Ala Glu Glu Asp Pro Glu Val His Leu Gly Gln Leu Lys
275 280 285

Arg Phe Ser Leu Arg Glu Leu Gln Val Ala Ser Asp Gly Phe Ser Asn
290 295 300

Lys Asn Ile Leu Gly Arg Gly Gly Phe Gly Lys Val Tyr Lys Gly Arg
305 310 315 320

Leu Ala Asp Gly Thr Leu Val Ala Val Lys Arg Leu Lys Glu Glu Arg
325 330 335

Thr Pro Gly Gly Glu Leu Gln Phe Gln Thr Glu Val Glu Met Ile Ser
340 345 350

Met Ala Val His Arg Asn Leu Leu Arg Leu Arg Gly Phe Cys Met Thr
355 360 365

Pro Thr Glu Arg Leu Leu Val Tyr Pro Tyr Met Ala Asn Gly Ser Val
370 375 380

Ala Ser Cys Leu Arg Glu Arg Pro Pro Ser Gln Pro Pro Leu Asp Trp
385 390 395 400

Pro Thr Arg Lys Arg Ile Ala Leu Gly Ser Ala Arg Gly Leu Ser Tyr
405 410 415

Leu His Asp His Cys Asp Pro Lys Ile Ile His Arg Asp Val Lys Ala

420 425 430

Ala Asn Ile Leu Leu Asp Glu Glu Phe Glu Ala Val Val Gly Asp Phe
435 440 445

Gly Leu Ala Lys Leu Met Asp Tyr Lys Asp Thr His Val Thr Thr Ala
450 455 460

Val Arg Gly Thr Ile Gly His Ile Ala Pro Glu Tyr Leu Ser Thr Gly
465 470 475 480

Lys Ser Ser Glu Lys Thr Asp Val Phe Gly Tyr Gly Ile Met Leu Leu
485 490 495

Glu Leu Ile Thr Gly Gln Arg Ala Phe Asp Leu Ala Arg Leu Ala Asn
500 505 510

Asp Asp Asp Val Met Leu Leu Asp Trp Val Lys Gly Leu Leu Lys Glu
515 520 525

Lys Lys Leu Glu Met Leu Val Asp Pro Asp Leu Gln Thr Asn Tyr Glu
530 535 540

Glu Arg Glu Leu Glu Gln Val Ile Gln Val Ala Leu Leu Cys Thr Gln
545 550 555 560

Gly Ser Pro Met Glu Arg Pro Lys Met Ser Glu Val Val Arg Met Leu
565 570 575

Glu Gly Asp Gly Leu Ala Glu Lys Trp Asp Glu Trp Gln Lys Val Glu
580 585 590

Ile Leu Arg Glu Glu Ile Asp Leu Ser Pro Asn Pro Asn Ser Asp Trp
595 600 605

Ile Leu Asp Ser Thr Tyr Asn Leu His Ala Val Glu Leu Ser Gly Pro
610 615 620

Arg
625

REFERENCES

Aleith, F. and Richter, G. (1990) Planta 183, 17-24.

Bent, A.F., Kunkel, B.N., Dahlbeck D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J. and Staskawicz, B.J. (1994) Science 265, 1856-1860.

Braun T, Schofield P.R. and Sprengel R. (1991) EMBO J. 10, 1885-1890.

Dangl, J.L. (1995) Cell 80, 363-366.

De Jong, A.J., Schmidt, E.D.L. and De Vries, S.C. (1993) Plant. Mol. Biol. 5, 367-377.

De Vries, S.C., Booij, H., Meyerink, P., Huisman, G., Wilde, H.D., Thomas T.L. and Van Kammen, A. (1988a) Planta 176, 196-204.

De Vries, S.C., Hoge, H., and Bisseling, T. (1988b) In Plant Molecular Biology Manual, B6, S.B. Gelvin, R.A. Schilperoort, and D.P.S. Verma. eds (Dordrecht, the Netherlands: Kluwer Academic Publishers), pp 1-13.

Dubois, T., Guedira, M., Dubois, D. and Vasseur, J. (1991) Protoplasma 162, 120-127.

Dudits, D., Györgyey, J., Bögre, L. and Bakó, L. (1995) In: In vitro Embryogenesis in Plants. Ed. Thorpe, T.A. Kluwer Press, pp. 267-308.

Engler J.A., Van Montagu M. and Engler G. (1994) Plant Mol. Biol. Rep. 12, 321-331.

Gamborg, O.L., Miller, R.A. and Ojima, K. (1968) Exp. Cell. Res. 50, 151-158.

Giorgetti, L., Vergara, M.R., Evangelista, M., LoSchiavo, F., Terzi, M. and Ronchi, V.N. (1995) Mol. Gen. Genet. 246, 657-662.

Goldberg, R.B., Barker, S.J. and Perez-Grau, L. (1989) Cell 56, 149-160.

Goldberg, R.B., Barker, S.J., Perez-Grau, L. (1989) Cell 56, 149-160.

Goldberg, R.B., de Paiva, G. and Yadegari, R. (1994) Science 266, 605-614.

Govind, S. and Steward, R. (1991) Dorsoventral pattern formation in *Drosophila*. Trends Genet. 7, 119-125.

Guzzo, F., Baldan, B., Levi, M., Sparvoli, E., LoSchiavo, F., Terzi, M., and P. Mariani (1995). Protoplasma 185, 28-36.

Guzzo, F., Baldan, B., Mariani, P., LoSchiavo, F. and Terzi, M. (1994) Exper. Botany 45, 1427-1432.

Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) Science 241, 42-52.

Hashimoto, C., Hudson, K.L. and Anderson K.V. (1988) Cell 52, 269-279.

Heck, G.R., Perry, S.E., Nichols, K.W. and Fernandez, D.E. (1995) AGL15, a MADS domain protein expressed in developing embryos. Plant Cell 7, 1271-1282.

Heldin, C-H. (1995) Cell 80, 213-223.

Hodge, R., Paul, Wyatt, Draper, J. and Scott, R. (1992) Plant J. 2, 257-260.

Horn, M.A. and Walker, J.C. (1994) Biochim. Biophys. Acta 1208, 65-74.

Jones, D.A., Thomas, C.M., Hammond-Kosack, K.E., Balint-Kurti, P.J. and Jones, J.D.G. (1994) Science 266, 789-792.

Kobe B. and Deisenhofer J. (1994) TIBS 19, 415-421.

Li, F., Barnathan, E.S. and Karikó, K. (1994) Nucl. Acid Res. 22, 1764-1765.

Liang, P. and Pardee, A.B. (1992) Science, 257, 967-971.

Meyerowitz, E.M. (1995) EDBC 95 congress of the European developmental biology organization. Abstract SI4.

Mu J-H., Lee H-S. and Kao T-h. (1994) Plant Cell 6, 709-721.

Pennell, R.I., Janniche, L., Scofield, G.N., Booij, H., de Vries, S.C. and Roberts, K. (1992) J. Cell Biol. 119, 1371-1380.

Rounsley, S.D., Ditta, G.S. and Yanofsky, M.F. (1995) Plant Cell 7, 1259-1269.

Sato, S., Toya, T., Kawahara, R., Whittier, R.F., Fukuda, H. and Komamine, A. (1995) Plant Mol. Biol. 28, 39-46.

Shelton, C.A. and Wasserman, S.A. (1993) Cell 72, 515-525.

Sterk, P., and De Vries, S.C. (1992) In Redenbaugh K (ed), Synseeds: Applications of synthetic seeds to crop improvement, CRC Press, London (1992).

Sterk, P., Booij, H., Schellekens, G.A., Van Kammen, A. and De Vries, S.C. (1991) Plant Cell 3, 907-921.

Thomas, T.L. (1993) Plant Cell 5, 1401-1410

Toonen, M.A.J. and De Vries, S.C. (1995) In: Embryogenesis, the generation of a plant. Eds. Wang, T.L. and Cuming, A. BIOS Scientific Publishers, Oxford, UK, pp. 173-189.

Toonen, M.A.J., Hendriks, T., Schmidt, E.D.L., Verhoeven, H.A., Van Kammen, A and De Vries, S.C. (1994) Planta 194, 565-572.

Toonen, M.A.J., Schmidt, E.D.L., Hendriks, T., Verhoeven, H.A., Van Kammen, A. and De Vries, S.C. (1996) *submitted.*

Torii, K.U. and Komeda, Y. (1994) 4th international congress of plant molecular biology. Abstract 692,

Van Engelen, F.A. and De Vries, S.G. (1992) Trends Genet. 8, 66-70.

Varner, J.E. and Lin, L.-S. (1989) Cell 56, 231-239.

Walker J.C. (1994) Plant Mol. Biol. 26, 1599-1609.

Walker, J.C. (1993) Plant J. 3, 451-456.

Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C. and Baker, B. (1994) Cell 78, 1101-1115.

Wilde, H.T., Nelson, W.S., Booij H., de Vries, S.C. and Thomas, T.L. (1988) Planta 176, 205-213.

Wurtele, E.S., Wang, H., Durgerian, S., Nikolau, B.J. and Ulrich, T.H. (1993) Plant Physiol. 102, 303-312.

Zhao, Y., Feng, X.H., Watson, J.C., Bottino, P.J. and Kung, S.D. (1994) Plant Molec. Biol. 26, 791-803.

Zimmerman, J.L. (1993) Plant Cell 5, 1411-1423.

What is Claimed is:

1. A method of producing apomictic seeds comprising the steps of:

   (i) transforming plant material with a nucleotide sequence encoding a protein the presence of which in an active form in a cell, or membrane thereof, renders said cell embryogenic,

   (ii) regenerating the thus transformed material into plants, or carpel-containing parts thereof, and

   (iii) expressing the sequence in the vicinity of the embryo sac.

2. A method according to the preceding claim, wherein the apomictic seeds are of the adventitious embryony type.

3. A method according to either of the preceding claims, wherein expression of the sequence yields a protein kinase capable of spanning a plant cell membrane.

4. A method according to the preceding claim wherein the kinase is capable of autophosphorylation.

5. A method according to any of the preceding claims, wherein the protein is a leucine rich repeat receptor like kinase and comprises a ligand binding domain, a proline box, a transmembrane domain, a kinase domain and a protein binding domain.

6. A method according to the preceding claim, wherein the protein lacks a functional ligand binding domain but comprises a proline box, a transmembrane domain, a kinase domain and a protein binding domain.

7. A method according to any preceding claim, wherein once incorporated into the cell membrane, the protein binding domain is located intra-cellularly.

8. A method according to any preceding claim, wherein the sequence further encodes a cell membrane targeting sequence.

9. A method according to any preceding claim, wherein the sequence is that depicted in SEQ ID Nos. 1 or 2 or is complementary to one which hybridizes under stringent conditions with the said sequences and which encodes a membrane bound protein having kinase activity.

10. A method according to any preceding claim, wherein the sequence is modified in that known mRNA instability motifs or polyadenylation signals are removed and/or codons which are preferred by the plant into which the sequence is to be inserted are used so that expression of the thus modified sequence in the said plant yields substantially similar protein to that obtained by expression of the unmodified sequence in the organism in which the protein is endogenous.

11. A method according to any preceding claim, wherein expression of the sequence is under control of an inducible or developmentally regulated promoter.

12. A method according to the preceding claim, wherein expression of the sequence is under control of one of the following: a promoter which regulates expression of SERK genes *in planta*, the carrot chitinase DcEP3-1 gene promoter, the *Arabidopsis* AtChitIV gene promoter, the *Arabidopsis* LTP-1 gene promoter, the *Arabidopsis* bel-1 gene promoter, the petunia fbp-7 gene promoter, the *Arabidopsis* ANT gene promoter, the promoter of the O126 gene from *Phalaenopsis*.

13. A method according to any of the preceding claims, wherein the sequence is expressed in the somatic cells of the embryo sac, ovary wall, nucellus, or integuments.

14. A method according to any of the preceding claims, wherein the endosperm within the apomictic seed results from fusion of polar nuclei within the embryo sac with a pollen-derived male gamete nucleus.

15. A method according to the preceding claim, wherein the sequence encoding the protein is expressed prior to fusion of the polar nuclei with the male gamete nucleus.

16. DNA comprising a sequence encoding a protein the presence of which in an active form in a cell, or membrane thereof, renders said cell embryogenic.

17. DNA according to claim 16, wherein the protein is a leucine rich repeat receptor like kinase and comprises a ligand binding domain, a proline box, a transmembrane domain, a kinase domain and a protein binding domain, the ligand binding domain optionally being absent or functionally inactive.

18. DNA according to either of claims 16 or 17 comprising a DNA sequence encoding a N-terminal protein fragment having the following amino acid sequence: Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val Thr Cys Asn.

19. DNA according to claim 18 comprising a DNA sequence encoding a N-terminal protein fragment having the following amino acid sequence: Val Xaa Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val Thr Cys Asn
with Xaa being a variable amino acid, but preferably Leu or Val.

20. DNA according to claim 19 comprising a DNA sequence encoding a N-terminal protein fragment having the following amino acid sequence: Val Xaa Gln Ser Trp Asp Pro Thr Leu Val Asn Pro Cys Thr Trp Phe His Val Thr Cys Asn Xab Xac Xad Xae Val Xaf Arg Val Asp Leu Gly Asn Xag Xah Leu Ser Gly His Leu Xai Pro Glu Leu Gly Xaj Leu Xak Xal Leu Gln
with Xaa to Xak being a variable amino acid, but preferably
Xaa = Leu or Val
Xab = Asn or Gln
Xac = Glu or Asp or His
Xad = Asn or His
Xae = Ser or Arg or Gln
Xaf = Ile or Thr
Xag = Ala or Ser
Xah = Glu or Asn
Xai = Val or Ala
Xaj = Val or Lys
Xak = Lys or Glu

Xal = Asn or His

21. DNA comprising a sequence encoding a protein having the sequence depicted in SEQ ID No. 3, or a protein substantially similar thereto which is capable of being membrane bound and which has kinase activity.

22. DNA comprising a sequence encoding a protein having the sequence depicted in SEQ ID No. 21, or a protein substantially similar thereto which is capable of being membrane bound and which has kinase activity.

23. DNA comprising a sequence encoding a protein having the sequence depicted in SEQ ID No. 33, or a protein substantially similar thereto which is capable of being membrane bound and which has kinase activity.

24. DNA comprising a sequence encoding a protein having the sequence depicted in SEQ ID Nos. 23, 25, 27, 29 and 31, or a protein substantially similar thereto which is capable of being membrane bound and which has kinase activity.

25. DNA according to any preceding claim, comprising a DNA having the sequence depicted in SEQ ID Nos. 1 or 2 or a sequence which is complementary to one which hybridizes under stringent conditions with the said sequences and which encodes a membrane bound protein having kinase activity.

26. DNA according to any preceding claim, comprising a DNA having the sequence depicted in SEQ ID No: 20 or a sequence which is complementary to one which hybridizes under stringent conditions with the said sequences and which encodes a membrane bound protein having kinase activity.

27. DNA according to any preceding claim, comprising a DNA having the sequence depicted in SEQ ID No: 32 or a sequence which is complementary to one which hybridizes under stringent conditions with the said sequences and which encodes a membrane bound protein having kinase activity.

28. DNA according to any preceding claim, comprising a DNA having the sequence depicted in SEQ ID Nos: 22, 24, 26, 28 and 30 or a sequence which is complementary to one which hybridizes under stringent conditions with the said sequences and which encodes a membrane bound protein having kinase activity.

29. DNA according to any of the preceding claims, which further encodes a cell membrane targeting sequence.

30. DNA according to any one of the preceding claims, in which the protein encoding region is under expression control of a developmentally regulated or inducible promoter.

31. DNA according to claim 30, wherein the promoter is one of the following: a promoter which regulates expression of SERK genes *in planta*, the carrot chitinase DcEP3-1 gene promoter, the *Arabidopsis* AtChitIV gene promoter, the *Arabidopsis* LTP-1 gene promoter, the *Arabidopsis* bel-1 gene promoter, the petunia fbp-7 gene promoter, the *Arabidopsis* ANT gene promoter, the promoter of the O126 gene from *Phalaenopsis*; the *Arabidopsis* DMC1 promoter, the pTA7001 inducible promoter.

32. DNA according to any preceding claim, wherein said DNA is a recombinant DNA.

33. DNA according to any preceding claim, wherein the sequence is modified in that known mRNA instability motifs or polyadenylation signals are removed and/or codons which are preferred by the plant into which the DNA is to be inserted are used so that expression of the thus modified DNA in the said plant yields substantially similar protein to that obtained by expression of the unmodified DNA in the organism in which the protein is endogenous.

34. DNA which is complementary to that which hybridizes under stringent conditions with the DNA of any one of claims 16 to 29.

35. A vector containing a DNA sequence as claimed in any one of claims 16 to 34.

36. Plant cell transformed with the DNA of any one of claims 16 to 34 or the vector of claim 35, which contains the DNA stably incorporated into its genome.

37. Plant cell according to claim 36, which is part of a whole plant.

38. Plants transformed with the DNA of any one of claims 16 to 34 or the vector of claim 35, the progeny of such plants which contain the DNA stably incorporated, and/or the apomictic seeds of such plants or such progeny.

39. Plants transformed with the DNA comprised by the recombinant DNA of claims 16 to 34.

40. Use of the DNA of any one of claims 16-34 in the manufacture of apomictic seeds.

41. Plants which are derived from apomictic seeds obtainable by the method of any one of claims 1-15 or 40.

42. A method of obtaining cultivars comprising the steps of fertilizing plants with the pollen of the plants of either of claims 38, 39 or 40 and cultivars which result from the said method.

43. A method of obtaining embryogenic cells in plant material, comprising transforming the material with a recombinant DNA sequence as claimed in any one of claims 16-34, the DNA comprised by the recombinant DNA of any one of claims 16 to 34, or the vector of claim 35, expressing the sequence in the material or derivatives thereof and subjecting the said material or derivatives to a compound which acts as a ligand for the gene product of the said sequence.

44. A method according to the preceding claim, wherein the sequence encodes a leucine rich repeat receptor like kinase, and the compound is a phyto-hormone.

45. A method of generating somatic embryos under *in vitro* conditions wherein the SERK protein is overexpressed ectopically.

46. A bag containing apomictic seeds obtainable by the method of any one of claims 1-15 or 40.

Fig. 1

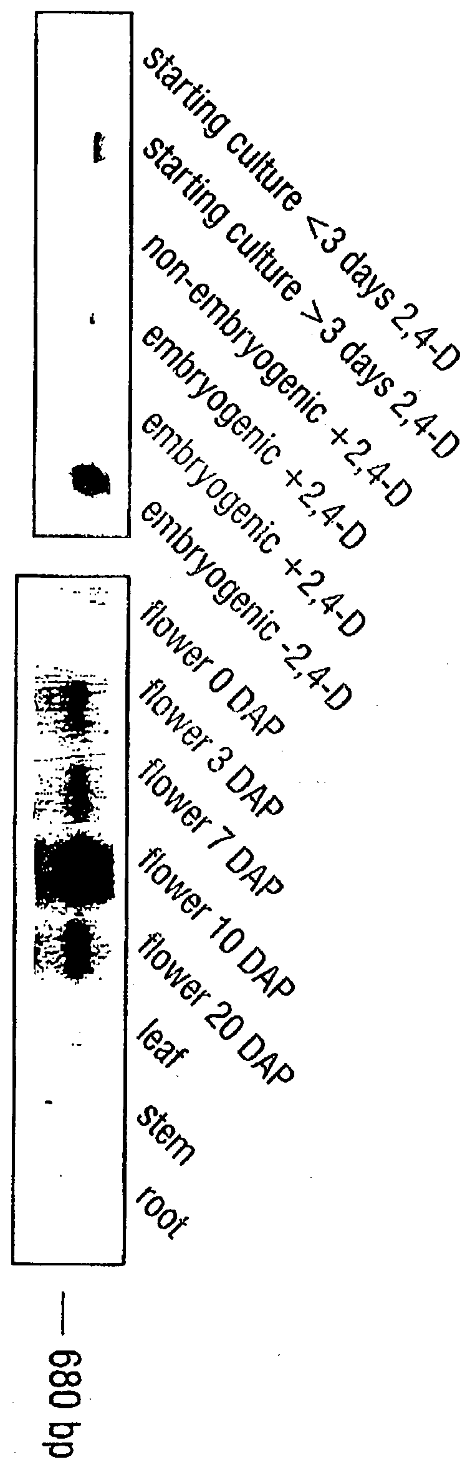

starting culture <3 days 2,4-D
starting culture >3 days 2,4-D
non-embryogenic +2,4-D
embryogenic +2,4-D
embryogenic +2,4-D
embryogenic -2,4-D
flower 0 DAP
flower 3 DAP
flower 7 DAP
flower 10 DAP
flower 20 DAP
leaf
stem
root
680 bp

Fig. 2

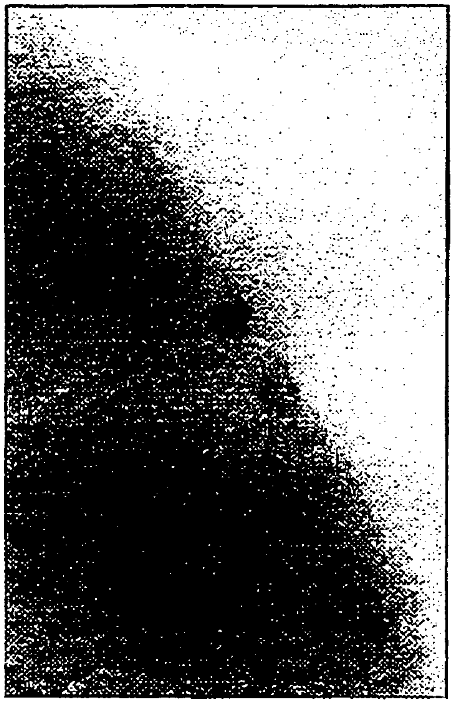


FIG. 2A

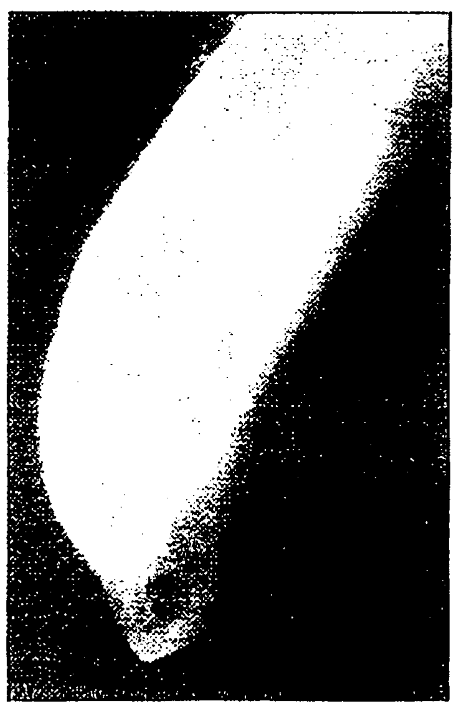


FIG. 2B

Fig. 3

FIG. 3 A

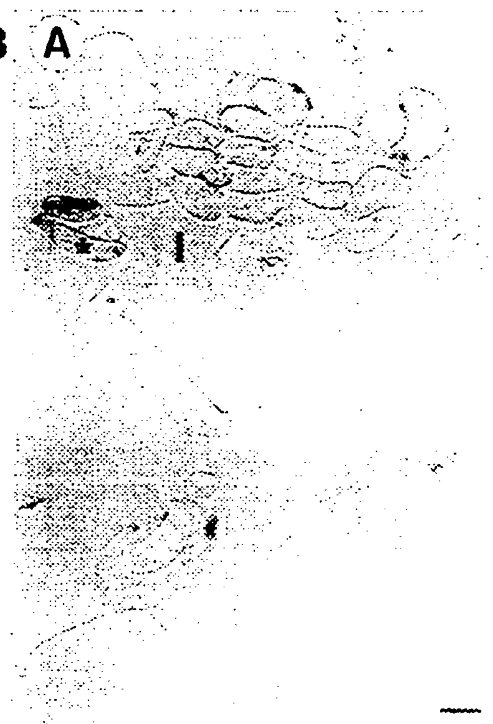


FIG. 3 B

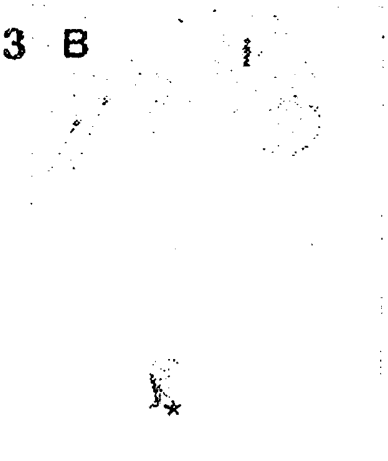


FIG. 3 C

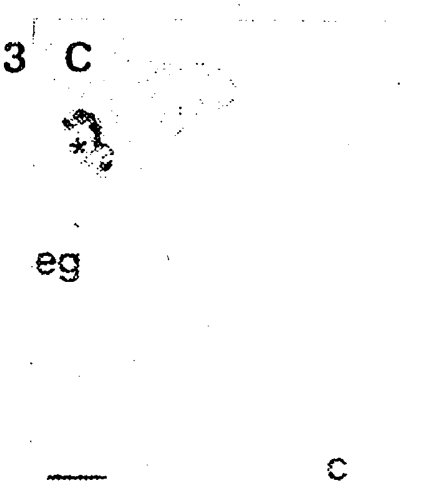


Fig. 3 (cont.)

FIG. 3 D

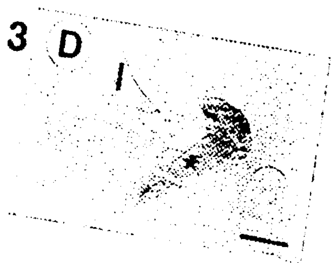


FIG. 3 E

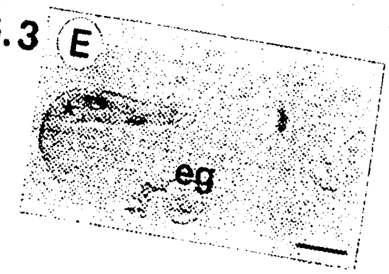


FIG. 3 F

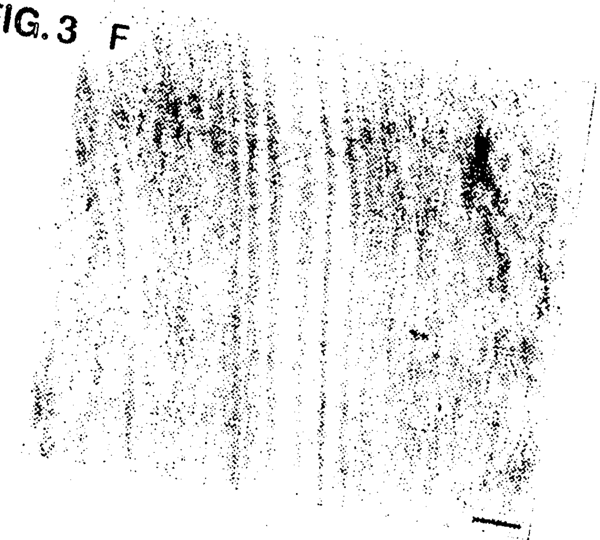


FIG. 3 G

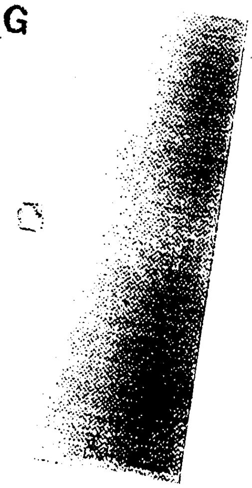


FIG. 3 H

Fig. 3 (cont.)

FIG. 3 I

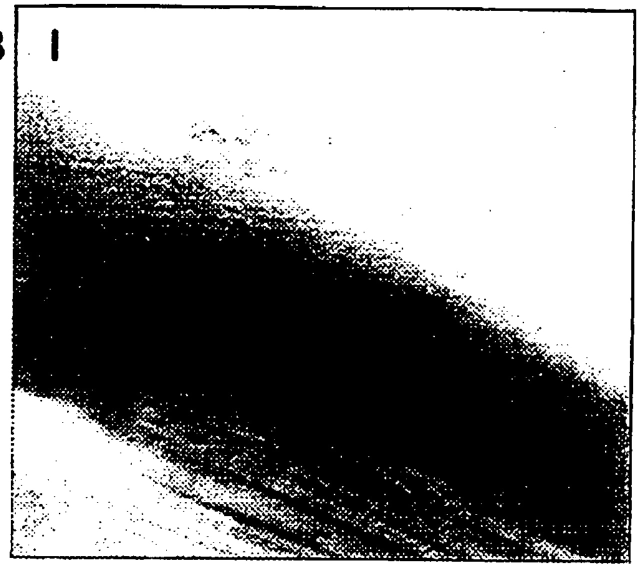


FIG. 3 H

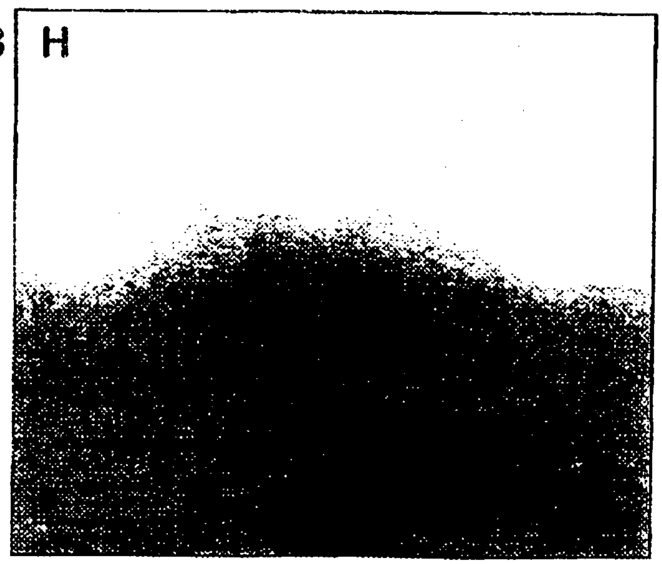


Fig. 4

no. I no. II no. III

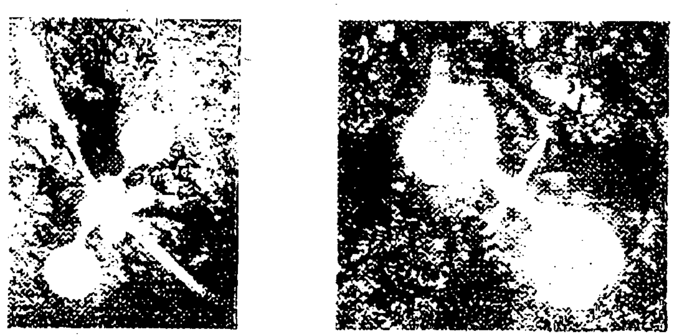


Fig. 5

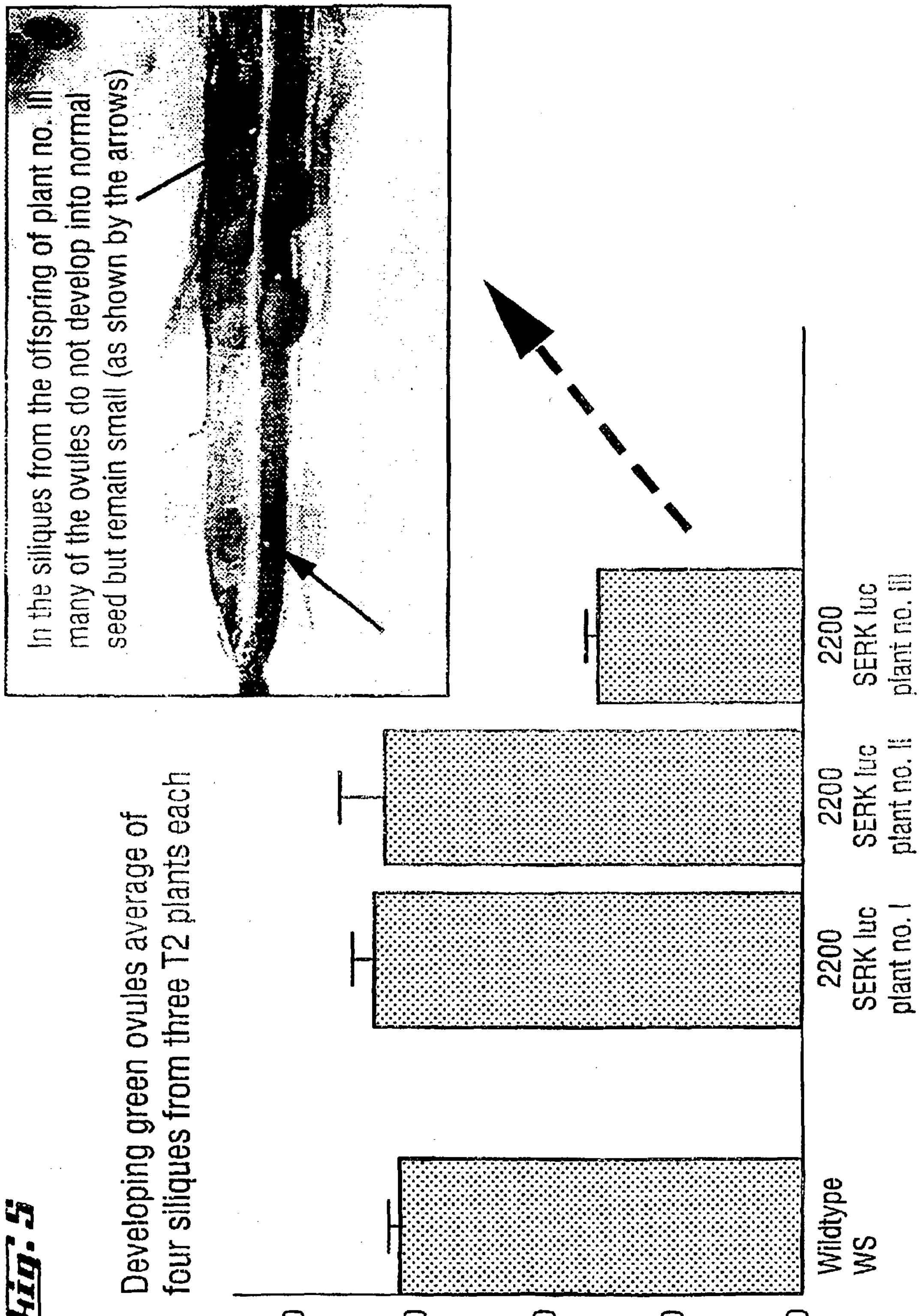
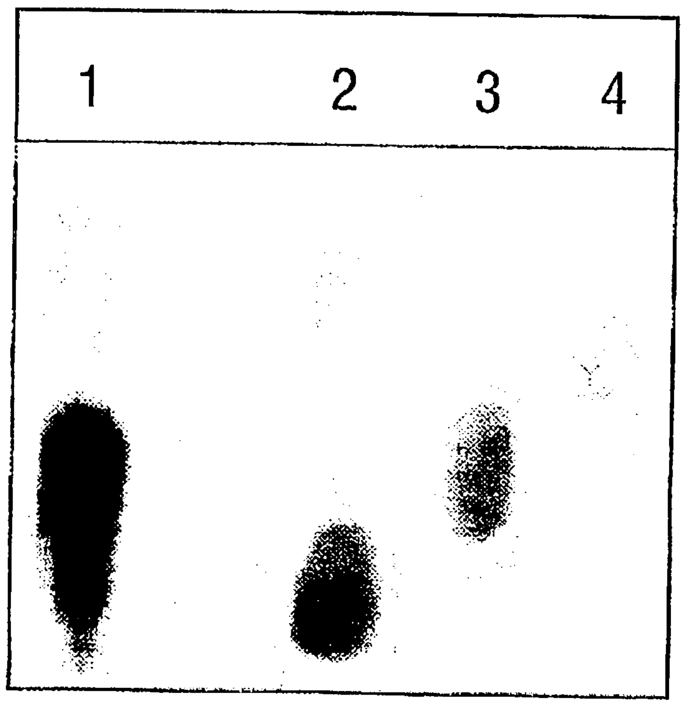

In the siliques from the offspring of plant no. III
many of the ovules do not develop into normal
seed but remain small (as shown by the arrows)
Developing green ovules average of
four siliques from three T2 plants each
40
30
20
10
0
Wildtype
WS
2200
SERK luc
plant no. I
2200
SERK luc
plant no. II
2200
SERK luc
plant no. III

Fig. 6



INTERNATIONAL SEARCH REPORT

Intern al Application No
PCT/EP 97/02443

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/82 C12N9/12 C07K14/415 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A01H C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HANNA, W.W. AND BASHAW, E.C.: "APOMIXIS: ITS IDENTIFICATION AND USE IN PLANT BREEDING" CROP SCIENCE, vol. 27, November 1987, pages 1136-1139, XP002040859 see page 1138, left-hand column ---	1,2

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
17 September 1997	30.09.97
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Holtorf, S

Form PCT/ISA/210 (second sheet) (July 1992)

WEST

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KOLTUNOW,A.M., ETAL . : "APOMIXIS: MOLECULAR STRATEGIES FOR THE GENERATION OF GENETICALLY IDENTICAL SEEDS WITHOUT FERTILIZATION" PLANT PHYSIOLOGY, vol. 108, 1995, pages 1345-1352, XP002040860	1,2,16, 35-41
Y	page 1345, right column, line 1-7; page 1347, right column; page 1348, right column, line 35-50; page 1349, 1350, 1351, left column, line 18-21; Fig. 2 + 3	9

X	KOLTUNOW, A.M.: "APOMIXIS: EMBRYO SACS AND EMBRYOS FORMED WITHOUT MEIOSIS OR FERTILIZATION IN OVULES" THE PLANT CELL, vol. 5, October 1993, pages 1425-1437, XP002040861	1,2,16, 35-41
Y	see the whole document	9

X	WO 89 00810 A (MAXELL HYBRIDS INC) 9 February 1989	1,2,16, 35-41
Y	pages 7,15,29-35	9

X	MARTIN, G.B.M., ET AL . : "A MEMBER OF THE TOMATO Pto GENE FAMILY CONFERS SENSITIVITY TO FENTHION RESULTING IN RAPID CELL DEATH" THE PLANT CELL, vol. 6, 1994, pages 1543-1552, XP002040862	21-28, 30,32, 34-39
Y	page 1546; Fig.4	9

X	SONG, W-Y, ET AL . : "A RECEPTOR KINASE-LIKE PROTEIN ENCODED BY THE RICE DISEASE RESISTANCE GENE , Xa21" SCIENCE, vol. 270, 15 December 1995, pages 1804-1806, XP002040863 see the whole document	24,30, 32,35-39

X	NEWMAN, T., ET AL . : "GENES GALORE: A SUMMARY OF METHODS FOR ACCESSING RESULTS FROM LARGE-SCALE PARTIAL SEQEUNCING OF ANONYMOUS ARABIDOPSIS cDNA CLONES" PLANT PHYSIOLOGY, vol. 106, 1994, pages 1241-1255, XP000571449 see the whole document	27,28

	-/--	

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NEWMAN,T., ET AL. : "GENES GALORE : A SUMMARY OF METHODS FOR ASSESSING RESULTS FROM LARGE-SCALE PARTIAL SEQEUNCING OF ANONYMOUS ARABIDOPSIS cDNA CLONES" EMBL SEQUENCE DATA LIBRARY, 30 August 1993, HEIDELBERG, GERMANY, XP002040864 ACCESSION No. T04109 ---	27,28
X	NEWMAN, T., ET AL . : "GENES GALORE: A SUMMARY OF METHODS FOR ASSESSING RESULTS FROM LARGE-SCALE PARTIAL SEQUENCING OF ANONYMOUS ARABIDOPSIS cDNA CLONES" EMBL SEQUENCE DATA LIBRARY, 27 July 1995, HEIDELBERG, GERMANY, XP002040865 ACCESSION No. H37195 ---	28
X	NEWMAN, T. , ET AL. : "GENES GALORE: A SUMMARY OF METHODS FOR ASSESSING RESULTS FROM LARGE-SCALE PARTIAL SEQUENCING OF ANONYMOUS ARABIDOPSIS cDNA CLONES" EMBL SEQUENCE DATA LIBRARY, 27 July 1995, HEIDELBERG, GERMANY, XP002040866 ACCESSION No. H36800 ---	28
X	DESPREZ, T. , ET AL . : "THE ARABIDOPSIS THALIANA TRANSCRIBED GENOME: THE GDR cDNA PROGRAM" EMBL SEQUENCE DATA LIBRARY, 25 June 1994, HEIDELBERG, GERMANY, XP002040867 ACCESSION No. Z34606 ---	28
A	DECROOCQ-FERRANT, V., ET AL.: "PETUNIA HYBRIDA HOMOLOGUES OF shaggy/zeste-white 3 EXPRESSED IN FEMALE AND MALE REPRODUCTIVE ORGANS" THE PLANT JOURNAL, vol. 7, no. 6, 1995, pages 897-911, XP002040868 see page 907, right-hand column ---	1-46
P,X	SCHMIDT, E.D.L., ET AL . : "A LEUCINE-RICH REPEAT CONTAINING RECEPTOR-LIKE KINASE MARKS SOMATIC PLANT CELLS COMPETENT TO FORM EMBRYOS" DEVELOPMENT, vol. 124, no. 10, 1997, pages 2049-2062, XP002040869 see the whole document ---	16-18, 21-32, 34-37, 39,45

-/--

1

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 97/02443

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 97 11167 A (US GOVERNMENT) 27 March 1997 page 1,5,21,24 ---	1,2,41
P,X	WO 97 10704 A (US GOVERNMENT) 27 March 1997 see the whole document -----	1,2,41

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

WEST

INTERNATIONAL SEARCH REPORT

Intern. al Application No
PCT/EP 97/02443

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 8900810 A	09-02-89	AU 629796 B	15-10-92
		AU 2255288 A	01-03-89
		EP 0329736 A	30-08-89
		CN 1040123 A	07-03-90
		JP 5503836 T	24-06-93
WO 9711167 A	27-03-97	AU 7367396 A	09-04-97
WO 9710704 A	27-03-97	AU 7367496 A	09-04-97